

IFW

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Appl. No. : 10/649,665 Confirmation No. 6920  
Inventor : N. NEMOTO et al.  
Filed : August 28, 2003  
Title : METHOD AND APPARATUS FOR MANAGING FAULTS  
IN STORAGE SYSTEM HAVING JOB MANAGEMENT  
FUNCTION  
Group Art Unit : 2165  
Examiner : S.G. Rimell  
Attorney Docket No.: ASA-1155  
Customer No. : 24956

Commissioner for Patents  
P.O. Box 1450  
Alexandria, VA 22313-1450

**TRANSMITTAL OF CERTIFIED PRIORITY DOCUMENT**

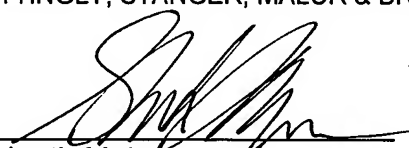
Sir:

Submitted herewith is a certified priority document (JP 2003-090518, filed March 28, 2003) of a corresponding Japanese patent application for the purpose of claiming foreign priority under 35 U.S.C. § 119. An indication that this document has been safely received would be appreciated.

Respectfully submitted,

MATTINGLY, STANGER, MALUR & BRUNDIDGE, P.C.

By

  
Shrinath Malur  
Reg. No. 34,663  
(703) 684-1120

October 11, 2005

# Best Available Copy

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日      2 0 0 3 年   3 月 2 8 日  
Date of Application:

出 願 番 号      特 願 2 0 0 3 - 0 9 0 5 1 8  
Application Number:

ST. 10/C] :      [ J P 2 0 0 3 - 0 9 0 5 1 8 ]

願      人      株式会社日立製作所  
Applicant(s):

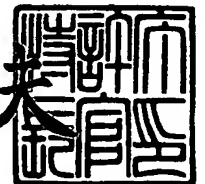
CERTIFIED COPY OF  
PRIORITY DOCUMENT

USSN 10/649,665  
MATTINGLY, STANGER, MALUR & BRUNDIDGE  
(703) 684-1120  
DKT: ASA-1155

2 0 0 3 年   8 月 2 6 日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康 夫



出証番号   出証特 2 0 0 3 - 3 0 6 9 6 3 2

【書類名】 特許願

【整理番号】 NT03P0246

【提出日】 平成15年 3月28日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 12/00

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 根本 直一

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 茂木 和彦

【発明者】

【住所又は居所】 神奈川県川崎市麻生区王禅寺 1 0 9 9 番地 株式会社日立製作所 システム開発研究所内

【氏名】 西川 記史

【発明者】

【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所 ソフトウェア事業部内

【氏名】 河村 信男

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100068504

【弁理士】

【氏名又は名称】 小川 勝男

【電話番号】 03-3661-0071

## 【選任した代理人】

【識別番号】 100086656

【弁理士】

【氏名又は名称】 田中 恭助

【電話番号】 03-3661-0071

## 【選任した代理人】

【識別番号】 100094352

【弁理士】

【氏名又は名称】 佐々木 孝

【電話番号】 03-3661-0071

## 【手数料の表示】

【予納台帳番号】 081423

【納付金額】 21,000円

## 【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 ジョブ管理機能を有するストレージ系障害管理方法及び装置

【特許請求の範囲】

【請求項 1】

ジョブを実行する計算装置と、複数の物理ディスク装置を収容するストレージ装置とを有するコンピュータシステムにおいて、特定の前記ジョブがアクセスするデータベース上の特定のテーブルに始まり、前記テーブルを格納するファイル、前記ファイルを格納する論理的なボリュームを経て前記ボリューム上のデータを分散して格納する前記物理ディスク装置に至るまでのデータマッピング情報であり、データマッピング経路上にあつて装置と他の装置とを接続するインタフェース部の識別子とを含むところの前記データマッピング情報に基づいて前記データマッピング経路上に存在する部位の障害によって影響を受けるジョブを特定する障害管理方法であつて、

障害を管理する計算機によって、前記ジョブ、前記テーブル、前記ファイル、前記ボリュームを経て前記物理ディスク装置に至るまでのデータマッピング上の対応情報を保持する装置の各々から前記データマッピング情報の一部情報を収集し、前記ジョブごとに前記データマッピング情報を集約して管理テーブルに格納し、前記ストレージ装置から前記物理ディスク装置のいずれかについて障害報告を受けたとき、前記管理テーブルを参照して障害によって影響を受けるジョブを特定して表示することを特徴とするジョブ管理機能を有するストレージ系障害管理方法。

【請求項 2】

前記データマッピング情報は、前記ファイルを格納する実ボリューム、前記実ボリュームを格納する 1 つ以上の仮想ボリューム、前記仮想ボリューム上のデータを分散して格納する論理ディスク装置および前記論理ディスク装置のデータを分散して格納する前記物理ディスク装置の階層構成を有し、

前記コンピュータシステムは、さらに前記計算装置と前記ストレージ装置との間に介在し、上位装置から受け取った入出力要求に含まれる前記仮想ボリュームの識別子を前記論理ディスク装置の識別子に変換する仮想化装置を有することを

特徴とする請求項 1 記載のジョブ管理機能を有するストレージ系障害管理方法。

【請求項 3】

前記データマッピング情報は、前記ファイルと前記実ボリュームとの対応関係、前記実ボリュームと前記仮想ボリュームとの対応関係、前記仮想ボリュームと前記論理ディスク装置との対応関係、および前記論理ディスク装置と前記物理ディスク装置との対応関係についての情報を含むことを特徴とする請求項 2 記載のジョブ管理機能を有するストレージ系障害管理方法。

【請求項 4】

ジョブがアクセスするデータベース上の特定のテーブルに始まり、前記テーブルを格納するファイル、前記ファイルを格納する論理的なボリュームを経て前記ボリューム上のデータを分散して格納する物理ディスク装置に至るまでのデータマッピング情報であり、データマッピング経路上にあつて装置と他の装置とを接続するインタフェース部の識別子とを含むところの前記データマッピング情報に基づいて前記データマッピング経路上に存在する部位の障害によって影響を受けるジョブを特定するコンピュータシステムであつて、

前記コンピュータシステムは、ジョブを実行する計算装置と、複数の前記物理ディスク装置を収容するストレージ装置と、障害を管理するサーバ計算機とを含み、

前記サーバ計算機は、前記ジョブ、前記テーブル、前記ファイル、前記ボリュームを経て前記物理ディスク装置に至るまでのデータマッピング上の対応情報を保持する装置の各々から前記データマッピング情報の一部情報を収集する手段と、前記ジョブごとに前記データマッピング情報を集約して管理テーブルに格納する手段と、前記ストレージ装置から前記物理ディスク装置のいずれかについて障害報告を受けたとき、前記管理テーブルを参照して障害によって影響を受けるジョブを特定して表示する手段とを有することを特徴とするコンピュータシステム。

【請求項 5】

前記データマッピング情報は、前記ファイルを格納する実ボリューム、前記実ボリュームを格納する 1 つ以上の仮想ボリューム、前記仮想ボリューム上のデー

タを分散して格納する論理ディスク装置および前記論理ディスク装置のデータを分散して格納する前記物理ディスク装置の階層構成を有し、

前記コンピュータシステムは、さらに前記計算装置と前記ストレージ装置との間に介在し、上位装置から受け取った入出力要求に含まれる前記仮想ボリュームの識別子を前記論理ディスク装置の識別子に変換する仮想化装置を有することを特徴とする請求項4記載のコンピュータシステム。

#### 【請求項6】

前記データマッピング情報は、前記ファイルと前記実ボリュームとの対応関係、前記実ボリュームと前記仮想ボリュームとの対応関係、前記仮想ボリュームと前記論理ディスク装置との対応関係、および前記論理ディスク装置と前記物理ディスク装置との対応関係についての情報を含むことを特徴とする請求項5記載のコンピュータシステム。

#### 【請求項7】

ジョブを実行する計算装置と、複数の物理ディスク装置を収容するストレージ装置とを有するコンピュータシステムにおいて、特定の前記ジョブがアクセスするデータベース上の特定のテーブルに始まり、前記テーブルを格納するファイル、前記ファイルを格納する論理的なボリュームを経て前記ボリューム上のデータを分散して格納する前記物理ディスク装置に至るまでのデータマッピング情報であり、データマッピング経路上にあつて装置と他の装置とを接続するインタフェース部の識別子とを含むところの前記データマッピング情報に基づいて前記データマッピング経路上に存在する部位の障害によって影響を受けるジョブを特定するサーバ計算機であつて、

前記ジョブ、前記テーブル、前記ファイル、前記ボリュームを経て前記物理ディスク装置に至るまでのデータマッピング上の対応情報を保持する装置の各々から前記データマッピング情報の一部情報を収集する手段と、前記ジョブごとに前記データマッピング情報を集約して管理テーブルに格納する手段と、前記ストレージ装置から前記物理ディスク装置のいずれかについて障害報告を受けたとき、前記管理テーブルを参照して障害によって影響を受けるジョブを特定して表示する手段とを有することを特徴とするサーバ計算機。

**【請求項 8】**

前記データマッピング情報は、前記ファイルを格納する実ボリューム、前記実ボリュームを格納する 1 つ以上の仮想ボリューム、前記仮想ボリューム上のデータを分散して格納する論理ディスク装置および前記論理ディスク装置のデータを分散して格納する前記物理ディスク装置の階層構成を有し、

前記コンピュータシステムは、さらに前記計算装置と前記ストレージ装置との間に介在し、上位装置から受け取った入出力要求に含まれる前記仮想ボリュームの識別子を前記論理ディスク装置の識別子に変換する仮想化装置を有することを特徴とする請求項 7 記載のサーバ計算機。

**【請求項 9】**

前記データマッピング情報は、前記ファイルと前記実ボリュームとの対応関係、前記実ボリュームと前記仮想ボリュームとの対応関係、前記仮想ボリュームと前記論理ディスク装置との対応関係、および前記論理ディスク装置と前記物理ディスク装置との対応関係についての情報を含むことを特徴とする請求項 8 記載のサーバ計算機。

**【請求項 10】**

ジョブを実行する計算装置と、複数の物理ディスク装置を収容するストレージ装置とを有するコンピュータシステムにおいて、特定の前記ジョブがアクセスするデータベース上の特定のテーブルに始まり、前記テーブルを格納するファイル、前記ファイルを格納する論理的なボリュームを経て前記ボリューム上のデータを分散して格納する前記物理ディスク装置に至るまでのデータマッピング情報であり、前記データマッピング情報は、前記ジョブ、前記テーブル、前記ファイル、前記ボリューム及び前記物理ディスクについての識別子間の対応関係と、データマッピング経路上にあつて装置と他の装置とを接続するインタフェース部の識別子とを含むところの前記データマッピング情報に基づいて前記データマッピング経路上に存在する部位の障害によって影響を受けるジョブを特定するサーバ計算機であつて、

前記データマッピング情報の一部情報を保持する装置の各々から前記一部情報を収集する手段と、前記ジョブごとに前記データマッピング情報を集約して管理



テーブルに格納する手段と、前記データマッピング経路上に存在する部位であってその識別子が前記管理テーブルに格納されている部位のいずれかについて障害報告を受けたとき、前記管理テーブルを参照して障害によって影響を受けるジョブを特定して表示する手段とを有することを特徴とするサーバ計算機。

【請求項 1 1】

前記データマッピング情報は、前記ファイルとこれを格納する実ボリュームとの対応関係、前記実ボリュームとこれを格納する 1 つ以上の仮想ボリュームとの対応関係、前記仮想ボリュームと前記仮想ボリューム上のデータを分散して格納する論理ディスク装置との対応関係、および前記論理ディスク装置と前記論理ディスク装置のデータを分散して格納する前記物理ディスク装置との対応関係についての情報を含み、

前記コンピュータシステムは、さらに前記計算装置と前記ストレージ装置との間に介在し、上位装置から受け取った入出力要求に含まれる前記仮想ボリュームの識別子を前記論理ディスク装置の識別子に変換する仮想化装置を有することを特徴とする請求項 1 0 記載のサーバ計算機。

【請求項 1 2】

前記サーバ計算機は、さらに前記障害によって影響を受けると特定された前記ジョブについて、あらかじめジョブ実行制御テーブルに設定された処理方法によって処理を行う手段を有することを特徴とする請求項 1 0 記載のサーバ計算機。

【請求項 1 3】

前記サーバ計算機は、さらにあらかじめ前記ジョブ実行制御テーブルに設定された管理者の確認の要否の情報に従い、管理者から承諾を指示されたとき前記処理方法によって処理を行う手段を有することを特徴とする請求項 1 2 記載のサーバ計算機。

【請求項 1 4】

ジョブを実行する計算装置と、複数の物理ディスク装置を収容するストレージ装置とを有するコンピュータシステムにおいて、特定の前記ジョブがアクセスするデータベース上の特定のテーブルに始まり、前記テーブルを格納するファイル、前記ファイルを格納する論理的なボリュームを経て前記ボリューム上のデータ

を分散して格納する前記物理ディスク装置に至るまでのデータマッピング情報であり、データマッピング経路上にあつて装置と他の装置とを接続するインタフェース部の識別子とを含むところの前記データマッピング情報に基づいて前記データマッピング経路上に存在する部位の障害によって影響を受けるジョブを特定するサーバ計算機で実行されるプログラムであつて、

前記サーバ計算機に、前記ジョブ、前記テーブル、前記ファイル、前記ボリュームを経て前記物理ディスク装置に至るまでのデータマッピング上の対応情報を保持する装置の各々から前記データマッピング情報の一部情報を収集する機能、前記ジョブごとに前記データマッピング情報を集約して管理テーブルに格納する機能、および前記ストレージ装置から前記物理ディスク装置のいずれかについて障害報告を受けたとき、前記管理テーブルを参照して障害によって影響を受けるジョブを特定して表示する機能を実現させるためのプログラム。

#### 【請求項 15】

前記データマッピング情報は、前記ファイルを格納する実ボリューム、前記実ボリュームを格納する 1 つ以上の仮想ボリューム、前記仮想ボリューム上のデータを分散して格納する論理ディスク装置および前記論理ディスク装置のデータを分散して格納する前記物理ディスク装置の階層構成を有し、

前記コンピュータシステムは、さらに前記計算装置と前記ストレージ装置との間に介在し、上位装置から受け取った入出力要求に含まれる前記仮想ボリュームの識別子を前記論理ディスク装置の識別子に変換する仮想化装置を有することを特徴とする請求項 14 記載のプログラム。

#### 【請求項 16】

前記データマッピング情報は、前記ファイルと前記実ボリュームとの対応関係、前記実ボリュームと前記仮想ボリュームとの対応関係、前記仮想ボリュームと前記論理ディスク装置との対応関係、および前記論理ディスク装置と前記物理ディスク装置との対応関係についての情報を含むことを特徴とする請求項 15 記載のプログラム。

#### 【請求項 17】

ジョブを実行する計算装置と、複数の物理ディスク装置を収容するストレージ

装置とを有するコンピュータシステムにおいて、特定の前記ジョブがアクセスするデータベース上の特定のテーブルに始まり、前記テーブルを格納するファイル、前記ファイルを格納する論理的なボリュームを経て前記ボリューム上のデータを分散して格納する前記物理ディスク装置に至るまでのデータマッピング情報であり、前記データマッピング情報は、前記ジョブ、前記テーブル、前記ファイル、前記ボリューム及び前記物理ディスクについての識別子間の対応関係と、データマッピング経路上にあって装置と他の装置とを接続するインタフェース部の識別子とを含むところの前記データマッピング情報に基づいて前記データマッピング経路上に存在する部位の障害によって影響を受けるジョブを特定するサーバ計算機で実行されるプログラムあって、

前記サーバ計算機に、前記データマッピング情報の一部情報を保持する装置の各々から前記一部情報を収集する機能、前記ジョブごとに前記データマッピング情報を集約して管理テーブルに格納する機能、および前記データマッピング経路上に存在する部位であってその識別子が前記管理テーブルに格納されている部位のいずれかについて障害報告を受けたとき、前記管理テーブルを参照して障害によって影響を受けるジョブを特定して表示する機能を実現させるためのプログラム。

#### 【請求項 18】

前記データマッピング情報は、前記ファイルとこれを格納する実ボリュームとの対応関係、前記実ボリュームとこれを格納する 1 つ以上の仮想ボリュームとの対応関係、前記仮想ボリュームと前記仮想ボリューム上のデータを分散して格納する論理ディスク装置との対応関係、および前記論理ディスク装置と前記論理ディスク装置のデータを分散して格納する前記物理ディスク装置との対応関係についての情報を含み、

前記コンピュータシステムは、さらに前記計算装置と前記ストレージ装置との間に介在し、上位装置から受け取った入出力要求に含まれる前記仮想ボリュームの識別子を前記論理ディスク装置の識別子に変換する仮想化装置を有することを特徴とする請求項 17 記載のプログラム。

#### 【請求項 19】

さらに前記サーバ計算機に、前記障害によって影響を受けると特定された前記ジョブについて、あらかじめジョブ実行制御テーブルに設定された処理方法によって処理を行う機能を実現させるための請求項 1 7 記載のプログラム。

【請求項 2 0】

さらに前記サーバ計算機に、さらにあらかじめ前記ジョブ実行制御テーブルに設定された管理者の確認の要否の情報に従い、管理者から承諾を指示されたとき前記処理方法によって処理を行う機能を実現させるための請求項 1 9 記載のプログラム。

【発明の詳細な説明】

【 0 0 0 1】

【発明の属する技術分野】

本発明は、計算機システムに用いられるストレージシステムに係わり、特にストレージ系障害の管理技術に関する。

【 0 0 0 2】

【従来の技術】

(1) SAN (Storage Area Network) について

近年、各サーバ計算機からストレージを切り離し、ストレージ系を集約化したストレージ入出力専用のネットワークである SAN が浸透しつつある。SAN の導入により、高速なデータ転送、ストレージシステムの高い拡張性と可用性およびストレージ資源の有効利用が実現可能となった。

(2) SAN 管理マネージャについて

SAN で集約されたストレージ装置を無停止で運用するためには、それぞれの機種の運用方法に精通した管理者が必要であり、管理コストが高くなる。ストレージ装置の運用管理においては、特に日々の運用の基本となる SAN に接続される各装置 (サーバ装置・仮想化スイッチ装置・ストレージ装置など) の稼動状況の監視が重要である。稼動状況を監視するためのソフトウェアを、以降、管理マネージャとよぶ。管理マネージャは、ストレージシステムの構成管理機能と障害監視機能の二大機能を持つ。構成管理機能とは、SAN を構成する各装置に存在する管理エージェントから情報を定期的 to 取得し、取得した情報から SAN の物理

的な接続関係(トポロジ)を検出し、常に最新のトポロジを視覚化して管理者に提供する機能である。障害監視機能とは、各装置が発行するハードウェア障害や性能低下などのイベント通知や、各装置に存在する管理エージェントから定期的に取得する装置情報に基づき、障害や性能低下等のイベントの発生を把握し、そのイベントを管理者に通知する機能である。これら二つの機能により、ユーザは、管理マネージャを用いて装置の稼動状況を一元的に管理でき、管理者の少人数化などによって運用コストの削減を図ることができる。

(3) データベース、ストレージ間のマッピング技術について

データベースとストレージ間のマッピング技術を使用して性能解析を行う技術が特許文献1に開示されている。このマッピング技術をSANに導入することにより、データベース又はファイル単位に性能解析を行うことができる。

【0003】

なおSANに関する技術として関連するものには、米国特許第6,035,306号などがある。

【0004】

【特許文献1】

米国特許第6,035,306号

【0005】

【発明が解決しようとする課題】

データベース管理システム(以下DBMSと呼ぶ)は、大規模なデータを格納するストレージ装置をデータベースとして使用する。DBMSの処理は、高いスループットと高信頼性が求められる。データベースを格納する記憶装置としてストレージ装置を使用することにより、高速かつ高信頼な処理を実行し続けることが可能である。ジョブにより呼び出されるDBMSは、ストレージ装置における複数の物理ディスク装置を意識することなく使用する。仮想化装置は、複数の論理ディスクを1つの仮想ボリュームとして提供する。DBMSがマッピング技術を使用して、バーチャリゼーション機能を持つ仮想化装置を含めたストレージ装置の構成を把握し、その管理を行うことは困難である。また仮想化装置を含むストレージ系に障害が発生したとき、特定の装置についての障害情報だけではどの

ジョブに影響を与えるものであるのか区別することができない。このような障害に際しては、人手による調査が必要である。また障害によって影響を受けるジョブの実行制御をする処理は、人手を要するものであり、高コストの要因となっている。

#### 【0006】

本発明の目的は、ストレージ系のいずれかの部位に障害が発生したとき、この障害によって影響を受けるジョブを特定し、これらジョブの実行制御をする技術を提供することにある。

#### 【0007】

##### 【課題を解決するための手段】

本発明は、特定のジョブがアクセスするデータベース上の特定のテーブルに始まり、このテーブルを格納するファイル、このファイルを格納する論理的なボリュームを経てそのボリューム上のデータを分散して格納する物理ディスク装置に至るまでのデータマッピング情報に基づいてデータマッピング経路上に存在するいずれかの部位の障害によって影響を受けるジョブを特定する障害管理技術を特徴とする。

#### 【0008】

また本発明は、障害によって影響を受けると特定されたジョブについて、あらかじめジョブ実行制御テーブルに設定された処理方法に従って処理を行う技術を特徴とする。

#### 【0009】

##### 【発明の実施の形態】

以下に、図面を参照しながら本発明の実施形態について説明する。なおこれにより本発明が限定されるものではない。

##### (1) 実施例1：各装置間のボリュームマッピングの構成管理と障害監視

本実施形態のコンピュータシステムは、管理サーバ装置が、仮想化スイッチ装置、DBMSサーバ装置及びストレージ装置と専用の管理ネットワークにより接続されている。DBMSサーバ装置、仮想化スイッチ装置及びストレージ装置の各々は、管理エージェントを具備する。管理サーバ装置は、これらの管理エー

メントから各データマッピング情報を収集し、ジョブ単位でのデータマッピングの構成を管理する。また管理サーバ装置は、SNMP (Simple Network Management Protocol) の障害通知メッセージを用いて障害の発生した装置から受け取るメッセージとデータマッピング情報から、障害の発生によるジョブへの影響を管理者に対して通知する。

#### 【0010】

まずSANの構成例について説明する。図1から図5は、SAN及びSANに接続される各装置の構成例を示す。図6は、データマッピングの階層構成を示す。図7から図16は、DBMSサーバ装置、仮想化スイッチ装置及びストレージ装置の装置内に具備された管理情報を示す。

#### 【0011】

図1は、システムの構成例を示す。本実施形態のコンピュータシステムは、DBMSサーバ装置1010、仮想化スイッチ装置1020、ストレージ装置1030及び管理サーバ装置1000から構成される。1台以上のDBMSサーバ装置1010（以下DBMSサーバと略称する）は、DBMS1011、DBMSサーバ管理エージェント1012及びボリュームマネージャ1013を有する。1台以上の仮想化スイッチ装置1020（以下スイッチと略称する）は、スイッチ管理エージェント1021を有する。1台以上のストレージ装置1030は、ストレージ装置管理エージェント1031を有する。1台の管理サーバ装置1000（以下管理サーバと略称する）は、管理サーバ管理マネージャ1001と障害管理マネージャ1002を有する。またDBMSサーバ1010とは別に、管理用ネットワーク1050に接続されDBMSサーバ1010を利用するアプリケーションサーバあるいはクライアント端末を有してもかまわない。あるいはDBMSサーバ1010自体をアプリケーションサーバと呼んでもよい。図1のシステムは、DBMSサーバ1010、スイッチ1020およびストレージ装置1030がそれぞれ2台ずつ記載されている。しかし以降の説明は、その都合上、1台のDBMSサーバ1010、1台のスイッチ1020及び1台のストレージ装置1030がファイバチャネル1040（以下FCと略称する）を介して相互に接続されているものとして説明する。

#### 【0012】

DBMSサーバ1010は、ストレージ装置1030をデータベースとして使用するDBMS 1011、ボリュームを管理するボリュームマネージャ1013及びDBMSサーバ1010内のボリュームの構成を管理するDBMSサーバ管理エージェント1012を有する。スイッチ1020は、当該装置内の仮想ボリュームの構成を管理するスイッチ管理エージェント1021を有する。ストレージ装置1030は、スイッチ1020の仮想ボリュームとストレージ装置の物理ディスク1032間の構成を管理するストレージ装置管理エージェント1031を有している。

#### 【0013】

管理サーバ1000は、管理用ネットワーク1050を介してDBMSサーバ1010、スイッチ1020及びストレージ装置1030に接続されている。管理サーバ1000の管理サーバ管理マネージャ1001は、管理用ネットワーク1050を介してDBMSサーバ管理エージェント1012、スイッチ管理エージェント1021及びストレージ装置管理エージェント1031と通信できる。

#### 【0014】

管理サーバ管理マネージャ1001は、DBMSサーバ管理エージェント1012、スイッチ管理エージェント1021及びストレージ装置管理エージェント1031から各装置のマッピング情報を取得し、後述する一元的な管理処理を行うことにより、SANにおける仮想ボリューム、実ボリューム、論理ディスク及び物理ディスクの構成を管理する。

#### 【0015】

図2は、管理サーバ1000の構成例を示す。管理サーバ1000は、プロセッサ2002、主記憶2005、管理用ネットワーク1050に接続する管理I/F（インタフェース）2001、出力装置2003及び実トポロジリポジトリ2004を有する。これらの装置は、内部バス等の通信路2006を介して相互に接続される。出力装置2003は、管理サーバ管理マネージャ1001によって実行された処理の実行結果を出力する。実トポロジリポジトリ2004は、管理テーブルデータ等を格納する。主記憶2005には、管理サーバ管理マネージャ1001、障害管理マネージャ1002、管理サーバ管理テーブル2011及びイベント辞書2012が格納されている。管理サーバ管理マネージャ1001は、プロセッサ2002によって実行されるプログラムであり、各装置からのデータ



マッピング情報と障害通知メッセージを受け取る。障害管理マネージャ1002は、障害管理を行うプログラムである。管理サーバ管理テーブル2011は、データマッピング情報を保持するテーブルである。イベント辞書2012は、障害検出メッセージを解釈するための辞書情報を格納する。

#### 【 0 0 1 6 】

図 3 は、D B M S サーバ1010の構成例を示す図である。D B M S サーバ1010は、プロセッサ3002、主記憶3005、管理用ネットワーク1050に接続する管理 I / F 3001及びF C 1040に接続する一つ以上のデータ I / F 3004を有する。これらの装置は、内部バス等の通信路3006を介して相互に接続される。主記憶3005には、実行中のD B M S 1011、ボリュームを管理するボリュームマネージャ1013、D B M S サーバ I / F テーブル3012、D B M S サーバ管理テーブル3011及びD B M S サーバ管理エージェント1012が格納されている。D B M S サーバ I / F テーブル3012は、F C 1040への I / F を管理する。D B M S サーバ管理テーブル3011は、ジョブが操作しているD B M S 1011管理下のテーブルやボリュームマネージャ管理情報をまとめたテーブルである。D B M S サーバ管理エージェント1012は、D B M S サーバ管理テーブル3011を管理するプログラムである。

#### 【 0 0 1 7 】

ボリュームマネージャ1013は、スイッチ1020が提供する仮想ボリュームを認識し、1つ以上の仮想ボリュームをまとめて1つの仮想的な実ボリュームとしてD B M S サーバ1010に提供する。ボリュームマネージャ1013は、プロセッサ3002によって実行されるプログラムである。また、図 3 には記載されていないが、D B M S 1011は、ファイルシステムを経由してからボリュームマネージャ1013にアクセスする。ファイルシステムは、実ボリューム上のファイルを管理する。

#### 【 0 0 1 8 】

図 4 は、スイッチ1020の構成例を示す。スイッチ1020は、コントローラ4003、記憶領域4004、管理用ネットワーク1050に接続する管理 I / F 4001及びS A N に接続する複数のデータ I / F 4002を有する。記憶領域4004、管理 I / F 4001及びデータ I / F 4002は、コントローラ4003を介して相互に接続される。コントローラ4003は、S A N のF C 1040を介して送受信されるデータのスイッチング及び仮

想ストレージ機能を実現する。記憶領域4004には、スイッチ管理エージェント1021、ボリューム仮想化プログラム4011、F C 接続管理テーブル4012、スイッチ I / F テーブル4013及びスイッチ管理テーブル4014が格納されている。スイッチ管理エージェント1021は、管理サーバ管理マネージャ1001と通信してスイッチ1020の管理情報を送受信し、スイッチ管理テーブル4014を管理するためのプログラムである。ボリューム仮想化プログラム4011は、記憶領域のバーチャリゼーションを実現するためのプログラムである。F C 接続管理テーブル4012は、S A N を介したスイッチ1020、D B M S サーバ1010及びストレージ装置1030間の接続関係を示す情報を格納する。スイッチ I / F テーブル4013は、データ I / F 4002についての管理情報を保持する。スイッチ管理テーブル4014は、仮想ボリュームについての管理情報を保持する。

#### 【 0 0 1 9 】

ボリューム仮想化プログラム4011は、スイッチ管理テーブル4014を参照してD B M S サーバ1010から受け取った入出力要求に含まれる仮想ボリュームの識別子を論理ディスク装置の識別子に変換する。

#### 【 0 0 2 0 】

なお本実施例ではスイッチ1020は6個のデータ I / F 4002 (S1、S2、S3、S4、S5、S6) を有する構成としているが、データ I / F 4002の個数は1つ以上であればいくつでもよい。この例では識別子S1、S2、S3によって示されるデータ I / F 4002は、D B M S サーバ1010と接続される。また識別子S4、S5、S6によって示されるデータ I / F 4002は、ストレージ装置1030と接続される。

#### 【 0 0 2 1 】

図5は、ストレージ装置1030の詳細な構成例を示す。ストレージ装置1030は、ストレージ装置1030内の制御を行うコントローラ5003、記憶領域5004、管理用ネットワーク1050に接続する管理 I / F 5001、S A N の F C 1040に接続する一つ以上のデータ I / F 5002及び物理ディスク1032を有する。記憶領域5004、管理 I / F 5001及びデータ I / F 5002は、コントローラ5003を介して相互に接続される。物理ディスク1032は、D B M S サーバ1010およびスイッチ1020に提供する記憶領域を構成する1つ以上の物理的なディスク装置である。

## 【 0 0 2 2 】

記憶領域5004には、ストレージ装置管理エージェント1031、ストレージ装置 I / F テーブル5011及びストレージ装置管理テーブル5012が格納されている。ストレージ装置管理エージェント1031は、管理サーバ管理マネージャ1001と通信してストレージ装置1030の管理情報を送受信し、ストレージ装置管理テーブル5012を管理するためのプログラムである。ストレージ装置 I / F テーブル5011は、データ I / F 5002についての管理情報を保持する。ストレージ装置管理テーブル5012は、物理ディスク1032についての管理情報を保持する。

## 【 0 0 2 3 】

なお本実施例ではストレージ装置1030は2個のデータ I / F 5002 (T1、T2) と 1 1 個の物理ディスク1032 (SS1、SS2、. . . 、SS11) を有するが、データ I / F 5002及び物理ディスク1032の個数は1つ以上であれば幾つであってもよい。

## 【 0 0 2 4 】

図6は、データマッピングの階層構成の一例を示す。DBMSサーバ1010は2つのジョブ6010、それらのジョブ6010が操作する3つのテーブル6020、ファイルシステム6101上の2つのファイル6030及びボリュームマネージャ1013が管理する3つの実ボリューム6040とを有する。スイッチ1020は、3つの仮想ボリューム6050を有する。ストレージ装置1030は、5つの論理ディスク装置6060と11個の物理ディスク装置1032を有する。DBMSサーバ1010上のテーブルデータ6020は、ファイルシステム6101上のファイル6030に格納される。ファイル6030は、ボリュームマネージャ1013上の実ボリューム6040に分散して格納されている。その実ボリューム6040は、スイッチ1020の仮想ボリューム6050に格納される。1つの実ボリューム6040が複数の仮想ボリューム6050に分散して格納されてもよい。仮想ボリューム6050は、ストレージ装置1030の論理ディスク装置6060を介して物理ディスク装置1032へと分散して格納されている。

## 【 0 0 2 5 】

なお本実施例では、データマッピング全体のうち、主として実ボリューム6040以下の階層構成に注目したデータマッピングを特にボリュームマッピングと呼んでいる。

## 【0026】

図6に示すように、特定のジョブに始まり、論理的なボリュームを経て特定の物理装置に至るデータマッピングの経路を特定することができる。このデータマッピング経路上には、装置と装置とを接続する特定のインタフェース部又は特定の物理ディスクが存在する。

## 【0027】

図6の例では、ジョブ6010はDBMSサーバ1010内に存在しているが、図示しないアプリケーションサーバ、クライアント端末などの計算装置上に存在するジョブであってもよい。これらのジョブ6010は、所在する装置上で稼動するOS（オペレーティングシステム）及びそのジョブ管理機能の制御の下に実行される。なお本例のデータマッピングの階層構成は、より縮退された階層構成であってもよい。例えばDBMS 1011がなく、ジョブ6010が直接ファイル6030にアクセスするような構成でもよい。またスイッチ1020がなく、実ボリューム6040が直接論理ディスク装置6060にボリュームマッピングされるような構成でもよい。あるいはスイッチ1020及び論理ディスク装置6060がなく、実ボリューム6040が直接物理ディスク装置1032にボリュームマッピングされるような構成でもよい。

## 【0028】

図7は、DBMSサーバ1010が保持するDBMSサーバI/Fテーブル3012の一例を示す。DBMSサーバI/Fテーブル3012のDB-I/F欄7010は、SANへの接続I/F識別子を格納し、名前欄7020は各I/F識別子に与えられる名称を格納する。

## 【0029】

図8は、DBMSサーバ1010が保持するDBMSサーバ管理テーブル3011の一例を示す。本テーブルのDBMS-ID欄8010は、管理サーバ1000が一意に識別ができるDBMS識別子を格納する。ジョブID欄8020はジョブ識別子を格納する。テーブル名欄8030はDBMS 1010が操作しているテーブルの識別子を格納する。ファイル名欄8040はテーブル6020が格納されているファイル6030のファイル名を格納する。実ボリューム名欄8050は、そのファイル6030が格納されている実ボリューム6040の識別子を格納する。DB-I/F欄8060はDBMSサーバ1010

が保持する S A N 接続の I / F 識別子を格納する。S C S I - I D 欄 8070 は接続先の I / F 識別子を格納する。L U N 欄 8080 は S C S I ターゲットデバイス内のボリュームにアクセスするための論理装置番号 (L U N) の識別子を格納する。

#### 【 0 0 3 0 】

図 9 は、スイッチ 1020 が保持するスイッチ I / F テーブル 4013 の一例を示す。本テーブルのデータ I / F - I D 欄 9010 はスイッチ 1020 が保持する I / F の識別子を格納する。仮想データ I / F - I D 欄 9020 は、D B M S サーバ装置 1010 にデータ I / F 4002 の識別子として認識させている識別子を格納する。S C S I - I D 欄 9030 は、仮想ボリューム 6050 に割り当てた S C S I I D の識別子を格納する。

#### 【 0 0 3 1 】

図 1 0 は、スイッチ 1020 が保持する F C 接続管理テーブル 4012 の一例を示す。本テーブルのデータ I / F 欄 10010 はスイッチ 1020 が保持する S A N 接続 I / F の識別子を格納する。スイッチ側名前欄 10020 は、S A N 接続の各 I / F 4002 に割り当てられている名前を格納する。接続先名前欄 10030 は S A N 接続の各 I / F が接続されている接続先の名前を格納する。

#### 【 0 0 3 2 】

図 1 1 は、スイッチ 1020 が保持するスイッチ管理テーブル 4014 の一例を示す。まずスイッチ管理テーブル 4014 の仮想ボリューム欄 11100 について説明する。仮想データ I / F - I D 欄 11110 は仮想ボリューム I / F の識別子を格納する。S C S I - I D 欄 11120 は仮想ボリュームに割り当てられた S C S I 識別子を格納する。L U N 欄 11130 は仮想データ I / F を介して仮想ボリューム 6050 にアクセスするための S C S I の L U N を格納する。仮想ボリューム I D 欄 11140 は、仮想データ I / F を介してアクセスされる仮想ボリューム 6050 に対し任意に割り当てた識別子を格納する。

#### 【 0 0 3 3 】

次に論理ディスク欄 11200 について説明する。実データ I / F - I D 欄 11210 は、仮想ボリューム 6050 を構成している論理ディスク 6060 にアクセスするために用いられるスイッチ 1020 の I / F の識別子を格納する。S C S I - I D 欄 11220 は

、実データ I / F の接続先である S C S I ターゲットデバイスの識別番号を格納する。L U N 欄 11230 は、実データ I / F を介して論理ディスク 6060 にアクセスするとき用いられる S C S I の L U N を格納する。

#### 【 0 0 3 4 】

図 1 2 は、ストレージ装置 1030 が保持するストレージ装置 I / F テーブル 5011 の一例を示す。本テーブルのデータ I / F - I D 欄 12010 は、ストレージ装置 1030 が保持する S A N 接続の I / F の識別子を格納する。ストレージ側名前欄 12020 は S A N 接続の I / F に割り当てられている名前を格納する。仮想データ I / F I D 欄 12030 は、ストレージ装置 1030 上の物理ディスク 1032 にアクセスするために用いられるデータ I / F 5002 の識別子を格納する。S C S I I D 欄 12040 は、実データ I / F の接続先である S C S I ターゲットデバイスの L U N を格納する。

#### 【 0 0 3 5 】

図 1 3 は、ストレージ装置 1030 が保持するストレージ装置管理テーブル 5012 の一例を示す。まずストレージ装置管理テーブル 5012 の論理ディスク欄 13100 について説明する。仮想データ I / F I D 欄 13110 は、ストレージ装置 1030 上の物理ディスク 1032 にアクセスするために用いられるデータ I / F 5002 の識別子を格納する。S C S I - I D 欄 13120 は、実データ I / F の接続先である S C S I ターゲットデバイスの識別番号を格納する。L U N 欄 13130 は物理ディスク 1032 にアクセスするための S C S I の L U N を格納する。論理ディスク I D 欄 13140 は、仮想データ I / F を介してアクセスされる論理ディスク 6060 に対し任意に割り当てた識別子を格納する。

#### 【 0 0 3 6 】

もう一方の物理ディスク欄 13200 内の物理ディスク I D 欄 13210 は、ストレージ装置 1030 が保持する物理ディスクの識別子を格納する。S C S I I D 欄 13220 は、実データ I / F の接続先である S C S I ターゲットデバイスの識別番号を格納する。S C S I L U N 欄 13230 は、実データ I / F 5002 を介してアクセスされる論理ディスク 6060 について S C S I の L U N を格納する。

#### 【 0 0 3 7 】

図 1 4 は、管理サーバ1000が保持する管理サーバ管理テーブル2011の一例を示す。管理サーバ管理テーブル2011は、図 7 から図 1 3 に至る DBMS サーバ管理エージェント1012、スイッチ管理エージェント1021及びストレージ装置管理エージェント1031が管理する各テーブルの情報について後述する処理を施した結果を格納する。

#### 【 0 0 3 8 】

まず DBMS サーバ欄14100について説明する。DBMS - I D 欄14101は管理サーバ1000が一意に識別が出来る DBMS 識別子を格納する。ジョブ I D 欄14102はジョブ識別子を格納する。テーブル名欄14103は DBMS 1010が操作しているテーブルの識別子を格納する。ファイル名欄14104は、DBMS 1010が操作しているテーブル6020に対応してそのテーブル6020を格納するファイル6030のファイル名を格納する。ボリュームマネージャ欄14105は、そのファイル6030が格納されているボリュームの名称を格納する。DB - I / F 欄14105は、DBMS サーバ1010が保持する SAN 接続の I / F 識別子を格納する。

#### 【 0 0 3 9 】

次にスイッチ欄14200について説明する。スイッチ I / F 欄14201はスイッチ1020が保持する SAN 接続 I / F の識別子を格納する。仮想ボリューム欄14210内の仮想ボリューム I D 欄14211は仮想データ I / F を介してアクセスされる仮想ボリューム6050に対し任意に割り当てた識別子を格納する。仮想データ I / F - I D 欄14212は仮想ボリューム I / F の識別子を格納する。SCSI - I D 欄14213は仮想ボリュームに割り当てられた SCSI 識別子を格納する。LUN 欄14214は仮想データ I / F を介してアクセスされる仮想ボリューム6050についてその SCSI の LUN を格納する。

#### 【 0 0 4 0 】

次にストレージ装置実ボリューム欄14220について説明する。実データ I / F - I D 欄14221は、仮想ボリューム6050を構成している論理ディスク6060にアクセスするために用いられるスイッチ1020の I / F の識別子を格納する。SCSI - I D 欄14222は実データ I / F の接続先である SCSI ターゲットデバイスの識別番号を格納する。LUN 欄14223は実データ I / F を介してアクセスされる

論理ディスク装置6060についてその S C S I の L U N を格納する。

【 0 0 4 1 】

次にストレージ装置欄14300について説明する。データ I / F I D 欄14301はストレージ装置1030が保持する S A N 接続の I / F の識別子を格納する。論理ディスク欄14310内について説明する。仮想データ I / F I D 欄14311は、ストレージ装置1030上の物理ディスク1032にアクセスするために用いられるデータ I / F 5002の識別子を格納する。S C S I - I D 欄14312は、実データ I / F の接続先である S C S I ターゲットデバイスの識別番号を格納する。L U N 欄14313は物理ディスク1032にアクセスするための S C S I - L U N を格納する。論理ディスク I D 欄14314は、仮想データ I / F を介してアクセスされる論理ディスク6060に対し任意に割り当てた識別子を格納する。

【 0 0 4 2 】

次に物理ディスク欄14320について説明する。物理ディスク I D 欄14321はストレージ装置1030が保持する物理ディスクの識別子を格納する。S C S I I D 欄14322は、実データ I / F の接続先である S C S I ターゲットデバイスの識別番号を格納する。S C S I L U N 欄14323は実データ I / F 5002を介してアクセスされる論理ディスク装置6060についてその S C S I - L U N を格納する。

【 0 0 4 3 】

この図 1 4 に示した管理サーバ管理テーブル2011によって、ジョブ6010ごとに DBMS サーバ1010内のどのテーブル6020を参照し、どのファイル6030、ボリューム6040を使用して、またそれらがスイッチ1020においてどのような仮想ボリューム6050を経てストレージ装置1030上の物理ディスク1032にアクセスを行っているのかを把握することができる。本テーブルは、DBMS サーバ管理テーブル3011、スイッチ管理テーブル4014及びストレージ装置管理テーブル5012の変化の度に更新される。すなわち管理サーバ1000は、管理サーバ管理テーブル2011を通して常に各装置のデータマッピングに関する最新状態を監視しているものとする。

【 0 0 4 4 】

図 1 5 に管理サーバ1000が保持するイベント辞書2012の一例を示す。イベント辞書2012の各行は、障害コード15010に対して対応する障害／復旧部位15020、そ



の識別子15030及び障害ごとに付与される障害管理 I D 15040を有する。識別子15030は、障害が発生した装置より管理サーバ1000へ送信される障害通知メッセージ内のバリアブル・バインディングズ・フィールド16009の障害コードを示している。障害管理 I D 15040は、障害発生から障害復旧までその障害を管理するために用いられる識別子である。

#### 【 0 0 4 5 】

図 1 6 に、SNMP プロトコルの障害通知メッセージ（以下 T r a p メッセージと略称する）のフォーマットとそのメッセージの一例を示す。現在の S A N 管理ソフトによる障害監視機能は、T r a p メッセージを用いることが多く、本実施例においてはこの T r a p メッセージを用いて障害通知を行っている。

#### 【 0 0 4 6 】

図 1 6 （a）に示す SNMP プロトコルの T r a p メッセージフォーマットは、図示するフィールドから構成される。コミュニティ名16002は、メッセージ送信先を示す。P D U ( P r o t o c o l D a t a U n i t ) タイプ16003は、メッセージの種類を示す。エンタプライズ16004は送信元装置のベンダ名を示す。エージェント・アドレス16005は送信先 I P アドレスを示す。ジェネリック・トラップ・タイプ16006は、T r a p メッセージの種別を示す。タイムスタンプ16008はメッセージの送信時刻を示す。バリアブル・バインディングズのフィールド16009はメッセージ内容を格納する。

#### 【 0 0 4 7 】

P D U タイプ・フィールド16003の値が” 4 ”のとき、本メッセージは T r a p メッセージであると判断される。ジェネリック・トラップ・タイプフィールド16006の値が” 6 ”のとき、T r a p メッセージは送信元装置ベンダ固有の定義に基づく T r a p メッセージと判断される。このとき各ベンダによって定義されたスペシフィック・トラップ・タイプフィールド16007とバリアブル・バインディングズ・フィールド16009(図中下線部)の内容にもとづき T r a p メッセージを解釈する必要がある。

#### 【 0 0 4 8 】

図 1 6 （b）に示す T r a p メッセージは、ストレージ装置1030のハードウエ

ア障害を通知するために、管理サーバ1000の障害管理マネージャ1002に対して送信される T r a p メッセージの一例である。この例は、P D U タイプ・フィールド16103の値が” 4 ”、ジェネリック・トラップ・タイプフィールド16106の値が” 6 ”であるために本メッセージは T r a p メッセージであり、送信元装置ベンダ固有の定義に基づく T r a p メッセージであると判断される。

#### 【 0 0 4 9 】

またスペシフィック・トラップ・タイプ16107に障害種別、バリアブル・バイインディンクス16109に障害発生部位を示す障害コードを格納するよう管理者が定義している場合、本メッセージは、障害コード”30c1”の部位にハードウェア障害が生じている旨を示している。また本メッセージは、障害が発生した場合においてのみ障害発生装置から管理サーバ1000に対して通知するのではなく、障害の発生した装置において障害が復旧となった場合についても T r a p メッセージを用いて管理サーバ1000に通知されるものとする。その場合については、バリアブル・バイインディンクス16109に障害復旧部位を示す障害コードを格納する。

#### 【 0 0 5 0 】

図 1 7 は、管理サーバ1000において実行されるデータマッピング処理の処理手順を示すフローチャートである。この処理は、DBMSサーバ1010、スイッチ1020及びストレージ装置1030の各装置からの情報を用いて各ジョブ6010ごとにデータマッピングを実行する。管理サーバ管理マネージャ1001は、メモリ上に各装置の管理情報をコピーする領域を用意する（ステップ17010）。次に管理サーバ管理マネージャ1001は、管理ネットワーク1050を介してDBMSサーバ管理エージェント1012、スイッチ管理エージェント1021及びストレージ装置管理エージェント1031と通信を行い、図 8 から図 1 3 に至る各装置の管理情報をコピーする（ステップ17020）。次に管理サーバ管理マネージャ1001は、ステップ17020でコピーした各装置の管理情報を実トポロジリポジトリ2004に格納する（ステップ17030）。次に管理サーバ管理マネージャ1001は、ステップ17020でコピーした情報を基にして図 1 4 に示す管理サーバ1000独自の管理サーバ管理テーブル2011を作成する（ステップ17040）。次に管理サーバ管理マネージャ1001は、作成した管理サーバ管理テーブル2011を管理サーバ1000の保持する実トポロジリポジトリ2004

に格納する（ステップ17050）。

#### 【 0 0 5 1 】

DBMSサーバ1010、スイッチ1020及びストレージ装置1030の各装置から得られた図8から図13の管理情報は、各装置の状態により常に変化する。管理サーバ1000は、これらの管理情報が変化するごとに図17に示す処理手順に従って管理サーバ管理テーブル2011を更新する。

#### 【 0 0 5 2 】

図18は、管理サーバ管理テーブル2011を作成する処理ステップ17040の詳細を展開して示すフローチャートである。管理サーバ管理マネージャ1001は、管理ネットワーク1050を介してDBMSサーバ1010、スイッチ1020及びストレージ装置1030と通信を行い、各装置のDBMSサーバ管理テーブル3011、スイッチ管理テーブル4014及びストレージ装置管理テーブル5012を基にして管理サーバ管理テーブル2011を作成する。管理サーバ管理マネージャ1001は、作成する管理サーバ管理テーブル2011のすべてのエントリについて以下の処理を実行する。

#### 【 0 0 5 3 】

まず管理サーバ管理マネージャ1001は、メモリ上に作成する管理サーバ管理テーブル2011の領域を用意する（ステップ18010）。次に管理サーバ管理テーブル2011の各エントリについて以下の処理を実行する（ステップ18020）。DBMSサーバ欄14100の各エントリには、DBMSサーバ1010のDBMSサーバ管理エージェント1012との通信により得たDBMSサーバ管理テーブル3011のエントリをコピーし、登録する（ステップ18030）。登録されるエントリは、DBMS ID 8010、ジョブID 8020、テーブル名8030、ファイル名8040、実ボリューム名8050及びDB I/F 8060である。

#### 【 0 0 5 4 】

次に管理サーバ管理マネージャ1001は、DBMSサーバ管理テーブル3011のDB-I/F欄8060の値をキーとして、DBMSサーバI/Fテーブル3012を検索し、DB-I/F欄の値が一致する行の名前欄7020の識別子を検出する（ステップ18040）。次に検出した名前欄7020の識別子をキーとしてFC接続管理テーブル4012を検索し、名前欄7020の識別子が一致する行のデータI/F欄10010の識

別子を検出する。管理サーバ管理マネージャ1001は、管理サーバ管理テーブル2011内のスイッチ欄14200のスイッチ I / F 14201エントリに D B M S サーバ欄14100内の D B I / F 欄14106と対応するように検出したデータ I / F 欄の10010の識別子をコピーし、登録する（ステップ18050）。

#### 【 0 0 5 5 】

次にステップ18050で検出したデータ I / F 10010の識別子をキーとして、スイッチ I / F テーブル4013を検索し、データ I / F の識別子が一致する仮想データ I / F I D 9020と S C S I I D 9030を検出し、当該管理サーバ管理テーブル2011のスイッチ欄14200内の仮想ボリューム欄14210の仮想データ I / F I D 14212と S C S I I D 14213エントリへコピーする（ステップ18060）。

#### 【 0 0 5 6 】

次に管理サーバ管理マネージャ1001は、検出した仮想データ I / F I D 9020と S C S I I D 9030をキーとして、スイッチ管理テーブル4014を検索し、キー値の仮想データ I / F I D 9020と S C S I I D 9030と一致する仮想ボリューム欄11100内の仮想ボリューム I D 11140、論理ディスク欄11200内の実データ I / F I D 11210、S C S I I D 11220及び L U N 識別子11230を検出する。次に管理サーバ管理マネージャ1001は、検出した各々のエントリを管理サーバ管理テーブル2011のスイッチ欄14200内の仮想ボリューム欄14210の仮想ボリューム I D 14211および L U N 欄14214と、ストレージ装置実ボリューム欄14220内の実データ I / F I D 14221、S C S I I D 14222および L U N 欄14223の各欄へコピーする（ステップ18070）。

#### 【 0 0 5 7 】

次にステップ18070で検出した実データ I / F I D 11210をキーとして F C 接続テーブル4012を検索し、キー値と一致する接続先名前10030の識別子を抽出する（ステップ18080）。次にステップ18080で抽出した接続先名前10030の識別子をキーとしてストレージ装置 I / F テーブル5011を検索し、キー値と一致するストレージ装置1030のストレージ装置 I / F テーブル5011内のデータ I / F I D 12010エントリ、仮想データ I / F I D 12030エントリ及び S C S I I D 12040エントリを検出する。検出した各エントリを管理サーバ管理テーブル2011のストレ

ージ装置欄14300のデータ I / F I D エントリ 14301 および論理ディスク欄14310 内の仮想データ I / F I D 14311 エントリへコピーする（ステップ18090）。

#### 【 0 0 5 8 】

次に管理サーバ管理マネージャ1001は、ステップ18090にて抽出した仮想データ I / F I D 12030 と S C S I I D 12040 の各エントリをキーとしてストレージ装置管理テーブル5012を検索し、論理ディスク欄13100内の L U N 13130 および論理ディスク I D 13140 の各エントリと、物理ディスク欄13200内の物理ディスク I D 13210、S C S I I D 13220 および S C S I L U N 13230 の各エントリを検出する。管理サーバ管理マネージャ1001は、検出した各エントリを管理サーバ管理テーブル2011のストレージ装置欄14300の論理ディスク欄14310内の S C S I I D 14312、L U N 14313 および論理ディスク I D 14314 の各エントリと、物理ディスク欄14320内の物理ディスク I D 14321、S C S I I D 14322 および S C S I L U N 14323 の各エントリへコピーする（ステップ18100）。

#### 【 0 0 5 9 】

最後に管理サーバ管理マネージャ1001は、管理サーバ管理テーブル2011の全てのエントリが登録されているか判定し、登録完了していないエントリがある場合にはステップ18020に戻る（ステップ18110）。以上のステップにより管理サーバ管理テーブル2011作成処理が終了する。図 1 4 は、本処理により実行された管理サーバ管理テーブル2011の一部を示している。

#### 【 0 0 6 0 】

図 1 9 は、管理サーバ1000において実行される障害検出処理、及び管理者に対して障害により影響のあるジョブが存在する場合の内容通知処理を示すフローチャートである。管理サーバ1000は、障害発生装置からの T r a p メッセージを受信する（ステップ19010）。すなわち D B M S サーバ1010、スイッチ1020およびストレージ装置1030の各装置において発生した障害内容は、S N M P プロトコルの T r a p メッセージのバリアブル・バインディングズ16009内の障害コードとして管理サーバ1000へと伝えられる。障害管理マネージャ1002は、障害発生装置より送信された T r a p メッセージを受け取った後に、メッセージのバリアブル・バインディングズ16009内のコードをキーとして、自身が保持するイベント辞

書2012内を検索し、識別子15030を抽出し、どの装置の障害であるのか、またその障害部位を特定する処理を行う（ステップ19020）。障害部位は、管理サーバ管理テーブル2011に登録されているいずれかの装置のインタフェース部又は物理ディスク装置を含む。

#### 【0 0 6 1】

次に障害管理マネージャ1002は、特定した障害部位の識別子15030をキーとして管理サーバ管理テーブル2011を検索し、障害通知に必要な情報を抽出する（ステップ19030）。必要な情報は、障害部位によって異なる。抽出された情報は、キー値の障害管理 I D 15040を用いて抽出した情報と合わせて障害復旧まで保持される。図 2 0 は、このようにして抽出され保存される情報の一例を示す。

#### 【0 0 6 2】

ステップ19030で検出した結果が、実行中の各ジョブ6010に対して影響のある障害部位であるのか、そうでないのかについて、障害管理マネージャ1002は、検出した結果から判別を行う（ステップ19040）。影響があるジョブが存在すると判断された場合には、抽出した情報から表示に必要な情報を検出し、表示処理を行う（ステップ19050）。

#### 【0 0 6 3】

ステップ19030での抽出結果の一例を図 2 0 に示す。図 2 0 は、管理サーバ1000が T r a p メッセージより、イベント辞書2012の障害コード15010 “30c1” を抽出した場合の抽出結果であり、障害管理 I D 20050 “F3031” によって管理されている。この例では、ストレージ装置1030にのみ障害が発生したと想定しているため、図 2 0 は、障害部位が影響を与える D B M S サーバ1010と、ストレージ装置1030間における各 D B M S 1011内のジョブ6010ごとの関係を示している。ただしストレージ装置1030以外の装置に障害が発生した場合は、図 2 0 の形式と異なる。

#### 【0 0 6 4】

また各ジョブ6010に対して発生した障害が影響を及ぼさないとされた場合には、「実行中のジョブに対して影響はないが、障害が発生した」旨の表示を行うために必要な情報を、ステップ19030において抽出した情報から検出し、表示処理

を行う（ステップ19060）。

#### 【0065】

現在のSANを管理するソフトウェアによる障害監視機能は、IETF(Internet Engineering Task Force)で作成されたRFC1157「A Simple Network Management Protocol (SNMP)」によって定められたSNMPプロトコルのTrapメッセージを利用することが多い。しかしSANに接続されているDBMSサーバ1010、スイッチ1020、ストレージ装置1030の各装置間がスイッチ1020などで仮想ボリューム用いている場合がある。このために、各装置からのTrapメッセージのみでシステム全体のどの装置でどの部位に障害が発生し、どのような障害影響が発生するのか把握することは難しい。そこで本実施例は、管理サーバ1000に障害管理マネージャ1002を配し、管理サーバ管理テーブル2011と共に運用することによりこの問題を回避する。管理サーバ1000の障害管理マネージャ1002は、管理サーバ1000内に保持しているイベント辞書2012を用いて、各装置からのTrapメッセージの内容を解釈する。そして実トポロジリポジトリ2004に格納した管理サーバ管理テーブル2011情報とTrapメッセージの解釈情報を付きあわせて、ジョブ6010の実行に影響が出ると判断した場合には、障害管理マネージャ1002は、その旨について管理者に対して通知を行う。ここでTrapメッセージとは、各装置において障害が発生した場合に、管理サーバ1000に対して管理ネットワークを通して通知を行うメッセージを指している。

#### 【0066】

ここで、図16(b)に示す障害コード15010が“30c1”の場合を例に考える。この場合、ストレージ装置1030は、障害が発生した旨のTrapメッセージを管理サーバ1000に送信する。障害管理マネージャ1002は、Trapメッセージを受信した時に、イベント辞書2012を参照し、Trapメッセージ中の障害コード“30c1”をキーとしてイベント辞書2012内の障害コード行を検索する。そして検索した行の識別子15030を抽出する。次に障害管理マネージャ1002は、抽出した識別子15030をキーとして管理サーバ管理テーブル2011を検索する。これによってキーとなる識別子15030を含む行14500を抽出し、DBMSサーバと発生した障

害から影響を受けるジョブ6010を絞り込むことが可能となる。ここでは“Job1”というジョブ I D 14102で示されるジョブが影響を受けるものとされる。また抽出された情報から、ジョブ6010と障害発生装置との関係を図 2 0 内の行20500のように導く。そしてその導いた行20500内の D B M S - I D 20110、ジョブ I D 20120、D B I / F 20150、データ I / F I D 20210および障害発生部位である実ボリューム I D 20220の情報から、障害発生によりジョブ6010に影響が出る場合の表示処理を行う。

#### 【 0 0 6 7 】

この時の表示例を図 2 1 に示す。本表示例21000は、障害が発生した装置内のハードウェアの故障箇所と、障害の影響を受けるジョブとを共に表示する。これによって管理者は障害により受けるジョブ6010への影響を容易に把握することができる。

#### 【 0 0 6 8 】

また図 2 2 は、障害が発生したが、ジョブ6010には影響が及ばなかった場合の表示処理例を示す。表示例22000は、ジョブ6010が使用していないストレージ装置1030内の物理ディスク1032 “SS3” に障害が発生した場合の表示例22000を示す。

#### 【 0 0 6 9 】

上記表示例21000および22000は、一箇所の障害発生の例を示すが、複数の障害が発生した場合においても同様の表示により管理者に対して通知されるものとする。

#### 【 0 0 7 0 】

本実施例で示した D B M S サーバ1010、スイッチ1020及びストレージ装置1030の各装置間のボリュームマッピング構成の把握と、各装置における障害の発生監視を行うことにより、管理者は容易に実行中の各ジョブ6010について障害が及ぼす影響を把握することが可能となる。

#### ( 2 ) 実施例 2 : 障害発生時のジョブ実行制御

実施例 2 は、実施例 1 で示した障害発生した時に管理者に対して障害を通知する内容に加えて、各ジョブの実行順序をジョブの実行時に決めておくことにより



、障害の影響を受けるジョブに対する実行管理の技術を示す。以降の説明は、実施例1で示した障害箇所特定および管理者に対し障害通知に至るまでについては、同じステップを踏むため、変更点のみ説明する。

#### 【 0 0 7 1 】

図 2 3 は、図 2 に示した管理サーバ1000の主記憶2005に新たにジョブ実行制御テーブル24000を追加した管理サーバ1000の構成例を示す。ジョブ実行制御テーブル24000は、管理サーバ1000の実トポロジリポジトリ2004に格納されている。

#### 【 0 0 7 2 】

図 2 4 は、ジョブ実行制御テーブル24000のデータ構成の一例を示す。ジョブ実行制御テーブル24000のDBMS-ID欄24010は、DBMSサーバ装置1010内のDBMS 1011に割り振られた一意の識別子を格納する。ジョブID欄24020はDBMSサーバ装置1010に関して実行中のジョブ6010のジョブIDを格納する。実施装置欄24030はジョブ実行制御処理を行う装置を格納する。

#### 【 0 0 7 3 】

優先度欄24040はジョブの処理優先度を格納する。ジョブの優先度は、通常処理の優先度数値を“3”とし、最大（優先度が最も高い状態）で“5”となるように、ジョブ6010ごとの処理についての優先度が設定される。処理方法欄24050は事前に与えられた処理方法によって制御を行う固定方法と、ジョブ6010の実行時間を利用して制御を行う変動方法のどちらかを格納する。処理内容欄24060は、当該ジョブについて施されるジョブ実行制御処理の内容を格納する。本例の「閉塞処理」は、当該ジョブがアクセスしていたデータを閉塞する処理である。また「休止」は、当該ジョブを停止させることを意味する。

#### 【 0 0 7 4 】

確認処理欄24070は、実施装置欄24030に指定の装置について処理内容欄24060に指定の処理を実行する前に、管理者に指定処理の実行確認を行う必要があるか否かの区別を格納する。自動実施欄24080は確認処理欄24070に“不要”が設定された場合に指定する必要があり、処理内容欄24060内の処理を管理者に対して確認せずに自動実行を行うか否かについて登録する。

#### 【 0 0 7 5 】

ここで処理方法24050の登録する固定処理とは、事前に管理者によって登録されているジョブ実行制御を実施装置において実行する処理である。また変動処理とは、各装置より得られる I/O、CPU モニタから計算されるジョブ6010の実行時間の計算値に基づいて動的にジョブの実行制御を行う方法である。障害管理マネージャ1002は、計算されたジョブ6010の実行時間が障害発生前の実行時間よりも大きい場合に、優先度欄24040を参照する。実行ジョブ6010の優先度が通常処理優先度数値より低いならば、障害管理マネージャ1002は、処理内容欄24060に指定されている制御処理を行う。ジョブ実行制御テーブル24000の優先度24040、処理方法24050及び処理内容24060の各欄は、管理者によりジョブ1060の実行時に設定されるものとする。

#### 【 0 0 7 6 】

図 2 5 は、障害管理マネージャ1002が実行するジョブ実行制御処理の流れを示すフローチャートである。本処理は、特に指摘しない限り障害管理マネージャ1002が実行するものとする。まず障害管理マネージャ1002は、実トポロジリポジトリ2004から管理者により登録されたジョブ実行制御テーブル24000を取得する（ステップ25000）。管理者は、ジョブ6010の実行時にジョブ実行制御テーブル24000を作成するが、スケジューラを用いている場合には前もって登録しておく必要がある。次にある装置において障害が発生した場合には、実施例1と同様の手順によって障害が発生した装置から送信される T r a p メッセージより障害箇所の特定を行う（ステップ25010）。

#### 【 0 0 7 7 】

次に障害の影響を受けるジョブについてジョブ実行制御処理を実施するにあたって、管理者に対し実施の確認を求めるか否かについて、ジョブ実行制御テーブル24000の確認処理欄24070に指定される情報から判断を行う（ステップ25020）。ステップ25020で確認を求めると判断された場合にはステップ25030に移行し、確認を求めずにジョブ実行制御テーブル24000の指定に従ってジョブ実行制御処理を実施する場合にはステップ25031に移行する。

#### 【 0 0 7 8 】

次にステップ25020において確認を求めると判断された場合に、管理者に対し

て実施例1と同様の形で障害が発生した旨の表示を行う。その表示内容は、障害の影響を受けるジョブ6010へのジョブ実行制御を行うため管理者の確認を要求するという内容である。管理者が承諾した場合には、処理内容を選択するステップ25040に移行する。管理者が拒否した場合には、ジョブ実行制御処理を行わずに障害復旧まで待ち、ステップ25060に移行する（ステップ25030）。

#### 【0 0 7 9】

次にステップ25020でジョブ実行制御テーブル24000の情報に基づいて自動実行を行うと判断された場合に、管理者に対して実施例1と同様の形で障害が発生した旨の表示処理を行い、ジョブ実行制御テーブル24000の自動実施欄24080に指定される情報からジョブ実行制御処理を行うのか行わないのか判断をする（ステップ25031）。ステップ25031で自動実行を行うと判断された場合にはステップ25040に移行し、自動実行を行わないと判断された場合にはステップ25060へ移行する。

#### 【0 0 8 0】

ステップ25030で管理者が承諾した場合もしくはステップ25031で自動実行を行うと判断した場合、ジョブ実行制御方法について事前に決まっている処理を行うか、ジョブの負荷状況をみて優先度を変化させる処理を行うかについて判断する（ステップ25040）。その判断は、ジョブ実行制御テーブル24000の処理方法欄24050に指定されている情報に基づく。

#### 【0 0 8 1】

ステップ25040において固定処理と判断された場合、障害復旧に備えて現在実行している処理の状態を記憶装置に保存し、ジョブ実行制御テーブル24000の実施装置欄24030に指定される装置において、処理内容欄24060に指定される処理方法によって、処理を実行する（ステップ25041）。

#### 【0 0 8 2】

ステップ25040において変動処理と判断された場合、障害管理マネージャ1002は、以下の処理を実行する。まずDBMSサーバ1010、スイッチ1020およびストレージ装置1030の各装置についてI/O、CPUモニタの情報を取得する。なお管理サーバ1000は、各装置のI/O、CPUモニタ情報を監視しているものとす

る。I/Oは単位時間当り入出力されるデータ量、CPUデータは通常各装置のCPU稼働率の統計データが用いられる。各装置のI/O、CPUモニタ情報を監視しておくことにより、ジョブ実施装置のI/O、CPU負荷を見積ることができる。

#### 【0083】

各装置のモニタ情報と管理サーバ1000が管理している管理サーバ管理テーブル2011の情報を監視することによって、障害管理マネージャ1002は、各ジョブ6010と各装置の関係から、障害発生箇所から影響を受けるジョブ6010を検索できる。この検索結果によって、ジョブ6010の優先度を用いた柔軟な実行制御や、事前に決められた実行制御を行うことができる。これによって障害の影響を容易に把握し、障害の影響を受けるジョブ6010の明確にし、管理者によって指定するジョブの実行制御を行うことができる。

#### 【0084】

このI/O、CPUモニタ情報に基づいてジョブ実行時間の見積を計算することができるものとする。また障害前のジョブ実行時間を取得できるものとする。定型的なジョブであれば、通常のジョブ実行時間の統計データを利用することは容易である。障害管理マネージャ1002は、ジョブ実行時間の見積値が障害発生前のジョブ実行時間より上回るか否か判断を行う（ステップ25042）。

#### 【0085】

ステップ25042において上回ったと判断された場合には、そのジョブの実行終了が予定時刻を過ぎるおそれがある。この場合に、障害管理マネージャ1002は、当該ジョブについてジョブ実行制御テーブル24000の優先度欄24040を参照し、その優先度が通常処理の標準優先度数値“3”との比較を行う（ステップ25043）。ステップ25043において、該当ジョブの優先度が低い優先度である場合、現在実行している処理の状態を記憶装置に保存し、ジョブ実行制御テーブル24000の処理内容欄24060内の内容に従って処理を行う（ステップ25044）。逆に優先度が通常処理の標準優先度数値“3”よりも高い優先度または標準優先度数値“3”と同じである場合、実行制御処理を行わずにステップ25050に移行する。一方、ステップ25042において上回らないと判断された場合、実行制御処理を行わずに

ステップ25050に移行する。

#### 【 0 0 8 6 】

次に障害発生装置において障害復旧がなされたとき、障害発生装置より復旧した意味の障害コードを入れた T r a p メッセージが管理サーバ1000に対して送信される。管理サーバ1000は、その T r a p メッセージを受信した時、メッセージ中の障害コードをイベント辞書2012より検索し、復旧部位の識別子15030と障害管理 I D 15040を抽出する。

#### 【 0 0 8 7 】

障害管理マネージャ1002は、障害復旧時にステップ19030での抽出結果20000及び保存された処理状態を用いて実行中のジョブ実行制御処理からジョブ実行制御処理を行う前に実行していた処理に戻す（ステップ25050）。このために障害管理マネージャ1002は、抽出された識別子15030と障害管理 I D 15040をキーにして、抽出結果20000を検索し、D B M S サーバ欄20100内のジョブ I D 20120を抽出する。そのジョブ I D 20120について、障害管理 I D 20050で指示されたジョブ実行制御処理からジョブ実行制御処理を行う前に実行していた処理に戻すよう、障害管理マネージャ1002を通して該当装置に指示する（ステップ25050）。次に管理サーバ1000は、管理者に対して発生した障害が復旧した旨のメッセージを作成し、管理者に通知を行う（ステップ25060）。

#### 【 0 0 8 8 】

本実施例においては、ステップ25044にてジョブ実行制御テーブル24000に指示される実施装置について、現在実行している処理の状態を保存して、ジョブ実行制御テーブルに指示される処理を実行する。この場合に現在実行している処理状態の保存はどの装置で行ってもかまわない。

#### 【 0 0 8 9 】

ここでストレージ装置1030の物理ディスク1032 “SS1” で障害が発生し、事前に決められた処理である固定処理方法で閉塞処理を行うジョブ実行制御処理の場合24101を例にとって説明する。この場合、ストレージ装置1030の物理ディスク1032 “SS1” に障害が発生したという障害通知メッセージがストレージ装置1030から管理サーバ1000へ送信される。管理サーバ1000が行う障害部位の特定処理につ

いては、実施例1の通りである。

#### 【0090】

ステップ25000で管理サーバ1000の障害管理マネージャ1002がメッセージを受信したとき、メッセージ内の障害コード16109からイベント辞書2012と管理サーバ管理テーブル2011の参照により図20に示す抽出結果20000が得られる。障害管理マネージャ1002は、ステップ25030の管理者からのジョブ実行制御処理の実施許可を得た後に、ステップ25040からステップ25041においてジョブ実行制御テーブル24000の指示に従い設定された実施装置24030において、現在実行中の処理状態を保存し、指定された処理方法24050で指定された処理内容24060を実行する。

#### 【0091】

障害発生装置であるストレージ装置1030の物理ディスク1032において障害が復旧された場合には、ステップ25050で障害発生装置のストレージ装置1030から障害復旧コード“00a1”を含んだT r a pメッセージが管理サーバへ送信される。障害管理マネージャ1002は、このメッセージを受信したとき、メッセージ内の障害復旧コード“00a1”をキーとしてイベント辞書2012から識別子の“SS1”と障害管理IDの“F3031”情報15101を抽出し、障害発生時に作成した図20のテーブル20000から識別子が“SS1”であり、障害管理ID“F3031”で管理されているジョブ実行制御処理15101から、記憶していたジョブ実行制御処理を行う前に実行していた処理に戻すという処理を実行する。管理サーバ1000は、ステップ25060でストレージ装置1030で発生した障害が復旧し、ジョブ実行制御処理をから記憶してある処理に戻したというメッセージを管理者に対して表示する。

#### 【0092】

図26は、ストレージ装置1030の物理ディスク1032“SS1”に障害が発生し、影響をうけるジョブ6010“Job1”に対するジョブ実行制御処理実行の確認を管理者に求める通知例26000を示している。管理者に障害の発生箇所と影響を受けるジョブ6010、DBMSサーバ1010等を共に表示することによって、管理者は各ジョブ6010への影響を容易に把握することができる。このとき障害管理マネージャ1002は、管理者へジョブ実行制御処理を行う際の詳細で分かりやすい情報を提供

する。この通知例26000は、障害発生箇所が一箇所であるとしているが、複数箇所において障害が発生したとしても同様の表示を行うこととする。

#### 【0093】

また図27は、ストレージ装置1030において発生した物理ディスク1032“SS1”の障害復旧がなされた内容の管理サーバ1000から管理者への通知例27000を示している。

#### 【0094】

本実施例で示した障害を検知し、発生した障害への対応をジョブごとに、指定された方法の実行制御、または優先度を加味した実行制御、またはスケジュール処理という方法を用いることで、障害発生時におけるシステムのパフォーマンス低下中に対して、効率のよいシステム運用が可能となる。

#### 【0095】

なお以上で説明した実施例1および実施例2においては、仮想化スイッチ装置がバーチャリゼーション機能を備えているネットワーク構成を想定していたが、スイッチ機能を有する装置とは異なる装置をバーチャリゼーション装置として管理ネットワーク及びSANに接続した構成であっても、上述の実施例1および2は同様の処理により実現可能である。

#### 【0096】

##### 【発明の効果】

本発明により、ジョブごとにデータマッピング情報を集約することができ、この情報に基づいて障害によって影響を受けるジョブを特定することができる。また障害によって影響を受けたジョブの実行制御をすることができる。

##### 【図面の簡単な説明】

##### 【図1】

ストレージ系を含むコンピュータシステムの構成例を示す図である。

##### 【図2】

管理サーバ装置の構成例を示す図である。

##### 【図3】

DBMSサーバ装置の構成例を示す図である。

**【図 4】**

仮想化スイッチ装置の構成例を示す図である。

**【図 5】**

ストレージ装置の構成例を示す図である。

**【図 6】**

データマッピングの階層構成例を示す図である。

**【図 7】**

DBMS サーバ装置が保持する I/F テーブルの例を示す図である。

**【図 8】**

DBMS サーバ装置が保持する DBMS サーバ管理テーブルの例を示す図である。

**【図 9】**

仮想化スイッチ装置が保持する I/F テーブルの例を示す図である。

**【図 10】**

仮想化スイッチ装置が保持する F C 接続管理テーブルの例を示す図である。

**【図 11】**

仮想化スイッチ装置が保持するスイッチ管理テーブルの例を示す図である。

**【図 12】**

ストレージ装置が保持する I/F テーブルの例を示す図である。

**【図 13】**

ストレージ装置が保持するストレージ装置管理テーブルの例を示す図である。

**【図 14】**

管理サーバ装置が保持する管理サーバ管理テーブルの例を示す図である。

**【図 15】**

管理サーバ装置が保持するイベント辞書内にある解釈の一例を示す図である。

**【図 16】**

障害通知メッセージのフォーマットとそのメッセージの一例を示す図である。

**【図 17】**

管理サーバ装置によって実行される各装置の管理情報から管理サーバ管理テ



ブルを作成する処理の一例を示すフローチャートである。

【図 1 8】

管理サーバ装置において実行される管理サーバ管理テーブル作成の詳細処理内容の一例を示すフローチャートである。

【図 1 9】

管理サーバ装置において実行される障害箇所検出処理および通知処理の一例を示すフローチャートである。

【図 2 0】

影響のあるジョブの情報を抽出した結果の一例を示す図である。

【図 2 1】

障害発生によりジョブに影響が出た場合の表示処理例の一例を示す図である。

【図 2 2】

障害が発生したが、ジョブには影響がなかった場合の表示処理例の一例を示す図である。

【図 2 3】

ジョブ実行制御テーブルを追加した管理サーバ装置の構成例を示す図である。

【図 2 4】

管理サーバ装置が保持するジョブ実行制御テーブルの一例を示す図である。

【図 2 5】

管理サーバ装置が実行するジョブ実行制御処理の一例を示すフローチャートである。

【図 2 6】

ストレージ装置において障害が発生し、ジョブ I D “Job1” に対するジョブ実行制御処理実行の確認を管理者に求める通知の一例を示す図である。

【図 2 7】

ストレージ装置において発生した物理ディスク I D “SS1” の障害復旧がなされた旨の管理者への通知の一例を示す図である。

【符号の説明】

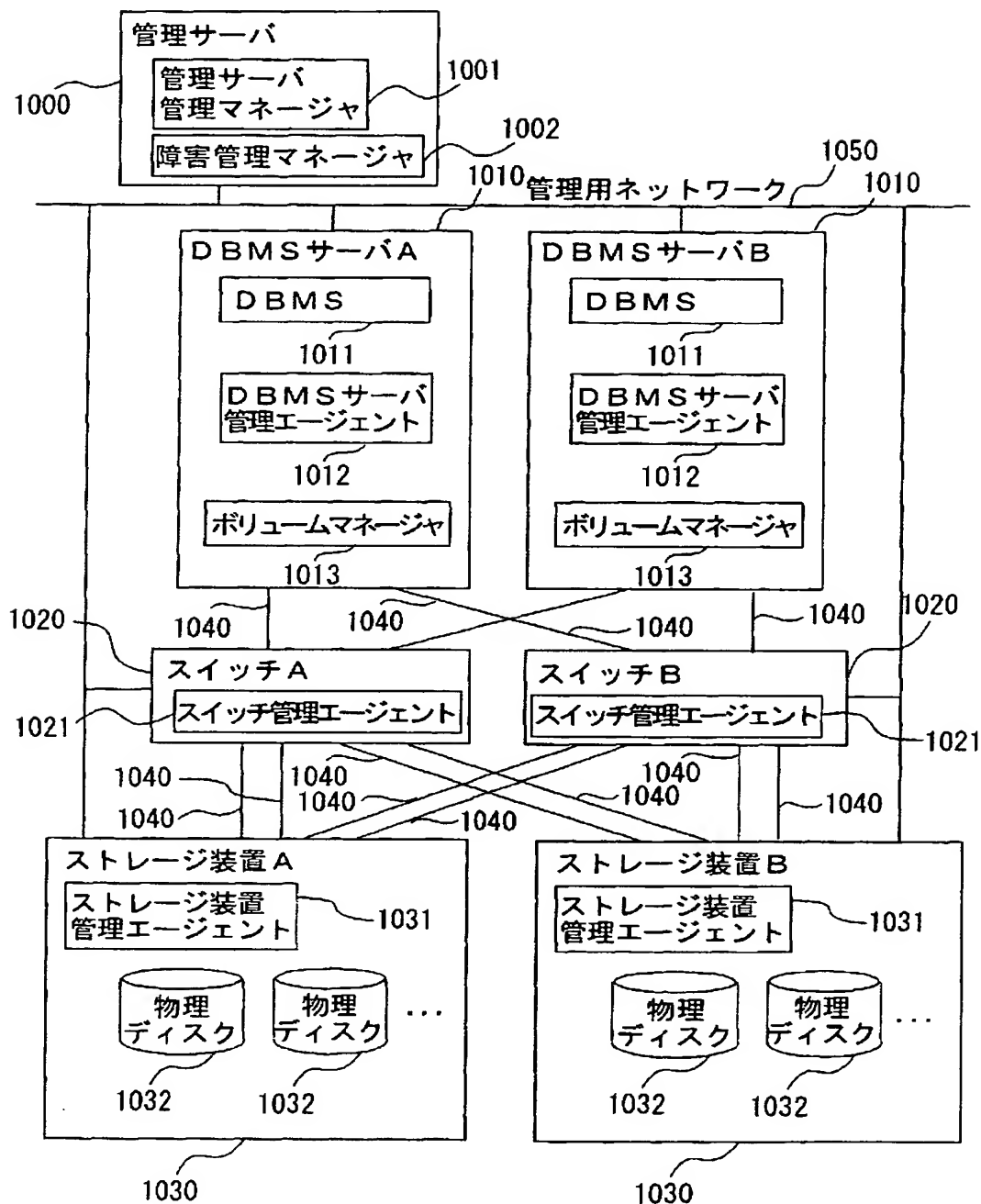
1000…管理サーバ装置、1010…D B M S サーバ装置、1020…仮想化スイッチ装

置、1030…ストレージ装置、1001…管理サーバ管理マネージャ、1002…管理サーバ障害管理マネージャ、1011…DBMS、1012…DBMSサーバ管理エージェント、1013…ボリュームマネージャ、1021…スイッチ管理エージェント、1031…ストレージ装置管理エージェント、1032…物理ディスク。

【書類名】 図面

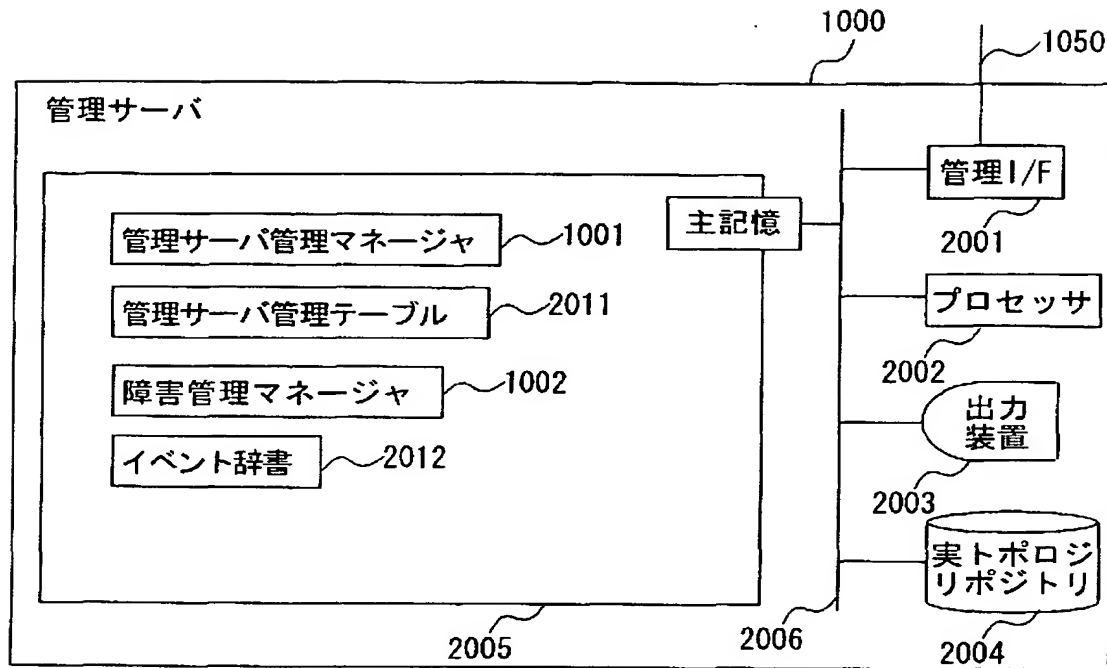
【図 1】

図 1



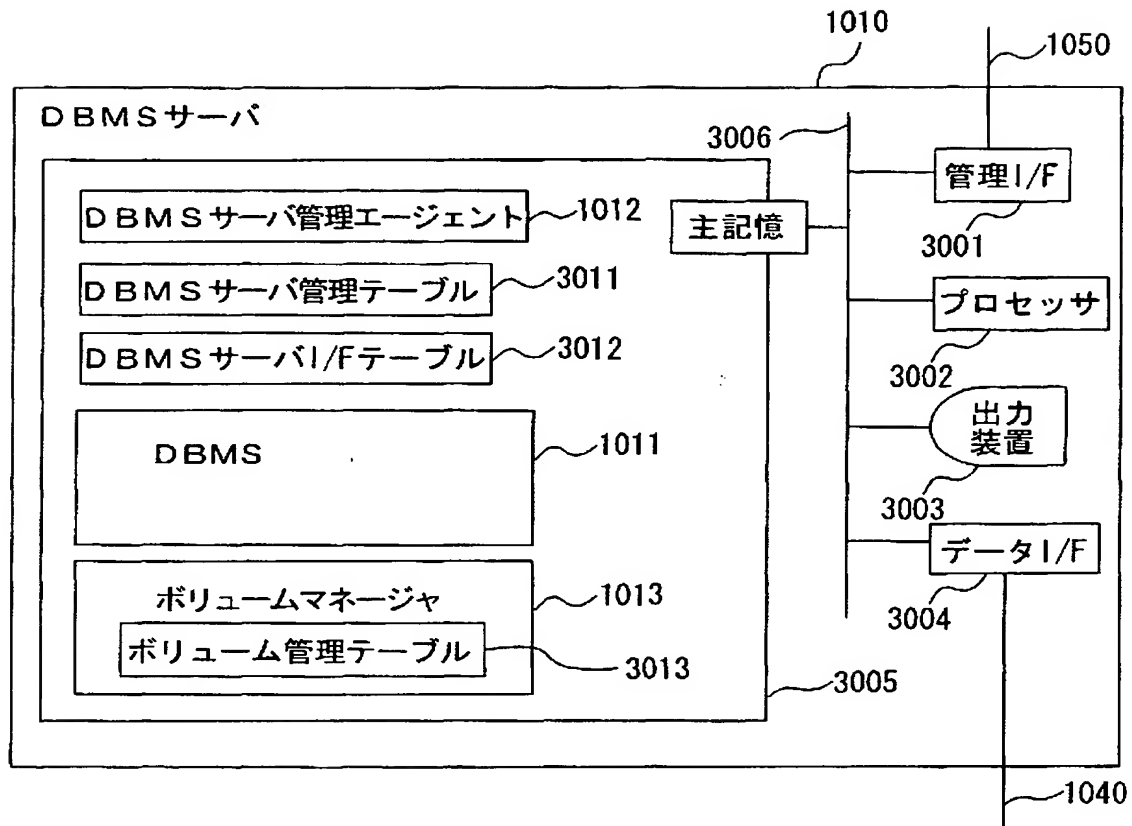
【図 2】

図 2



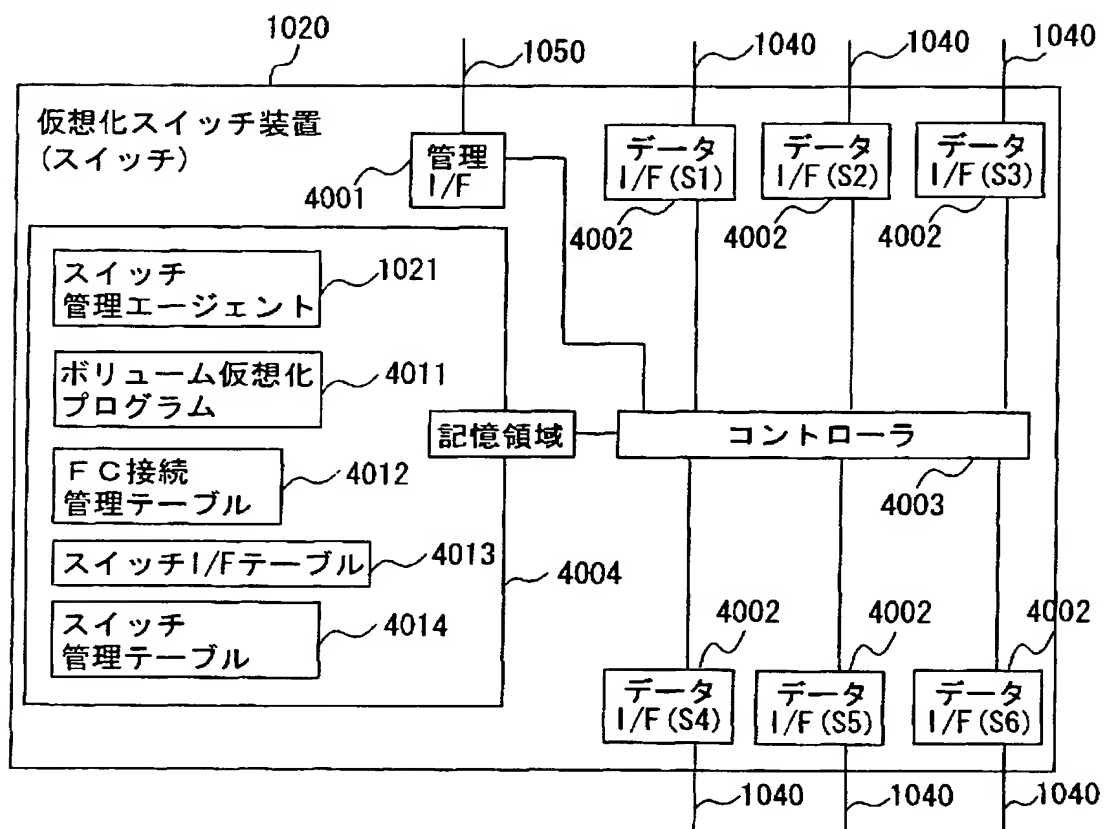
【図 3】

図 3



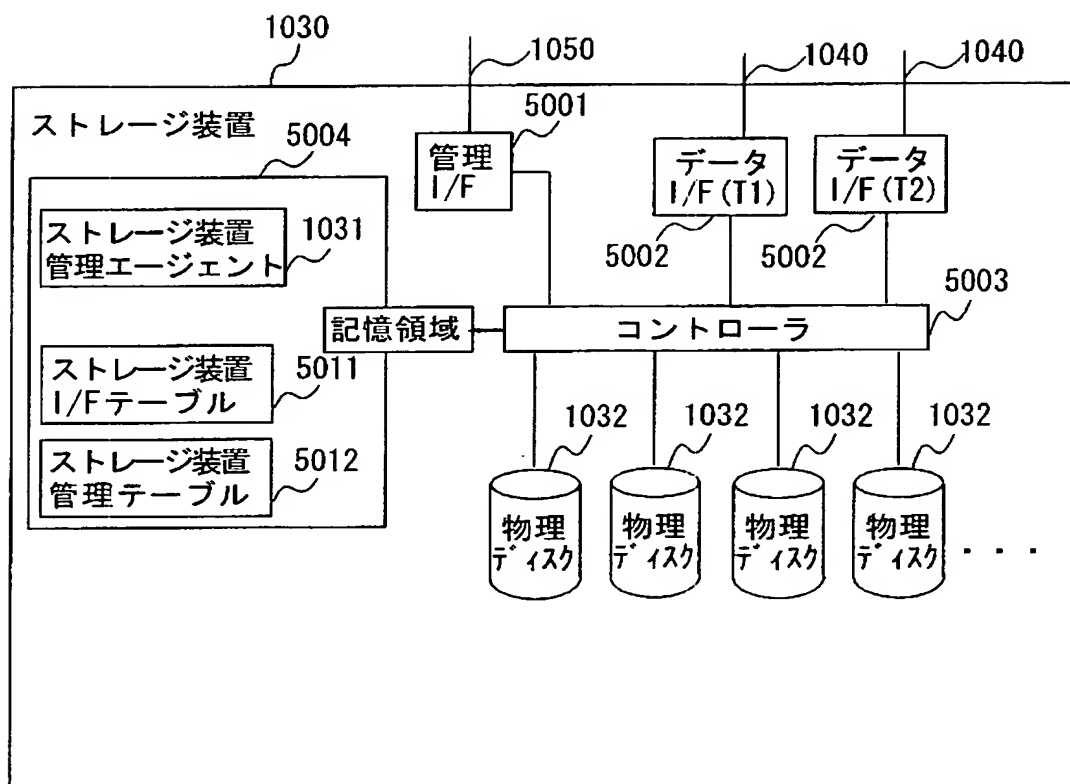
【図 4】

図 4



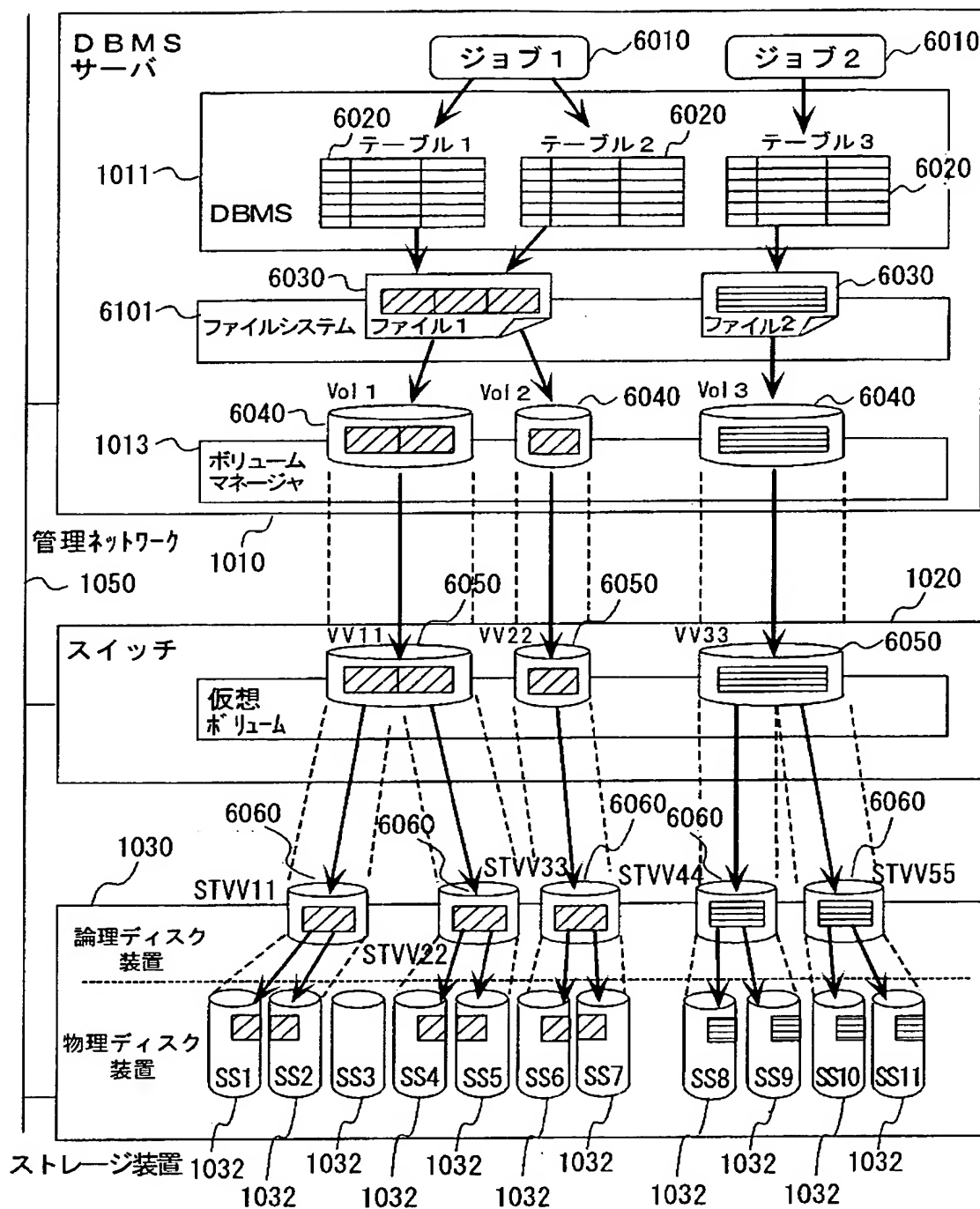
【図 5】

図 5



【図 6】

図 6





【図 7】

図 7

DB-I/F	名前
R 1	DB 1

3012 DBMSサーバI/Fテーブル

【図 8】

図 8

DBMS-ID	ジョブID	テーブル名	ファイル名	実ボリューム名	DB-I/F	SCSI-ID	LUN
DBMS1	Job1	表 1	File1	Vol1	R1	1	1
DBMS1	Job1	表 1	File1	Vol1	R1	2	1
DBMS1	Job1	表 2	File1	Vol2	R1	1	1
DBMS1	Job2	表 3	File2	Vol3	R1	1	1
DBMS1	Job2	表 3	File2	Vol3	R1	2	1

3011 DBMSサーバ管理テーブル

【図 9】

図 9

データ I/F-ID	仮想 データ I/F-ID	SCSI-ID
S1	LSWIF1	1
S1	LSWIF1	2
S2	LSWIF2	1
S3	LSWIF3	1
S3	LSWIF3	2

4013 スイッチI/Fテーブル

【図 10】

図 10

データ I/F	スイッチ 側名前	接続先名 前
S1	Switch1	DB1
S2	Switch2	DB1
S3	Switch3	DB1
S4	Switch4	STRAGE1
S5	Switch5	STRAGE1
S6	Switch6	STRAGE2

4012 FC接続管理テーブル

【図 11】

図 11

仮想ボリューム				論理ディスク		
仮想データ/F ID	SCSI-ID	LUN	仮想ボリュームID	実データ/F ID	SCSI-ID	LUN
LSWIF1	1	1	VV11	S4	1	1
LSWIF1	1	1	VV11	S4	1	2
LSWIF1	2	1	VV12	S4	2	1
LSWIF1	2	1	VV12	S4	2	2
LSWIF2	1	1	VV21	S5	3	1
LSWIF2	1	1	VV21	S5	3	2
LSWIF3	1	1	VV31	S6	1	1
LSWIF3	1	1	VV31	S6	1	2
LSWIF3	2	1	VV32	S6	2	1
LSWIF3	2	1	VV32	S6	2	2

4014 スイッチ管理テーブル

【図 12】

図 12

データ I/F ID	ストレージ側名前	仮想 データ I/F ID	SCSI-ID
T 1	STRAGE1	LSTIF1	1
T 1	STRAGE1	LSTIF2	2
T 1	STRAGE1	LSTIF3	3
T 2	STRAGE2	LSTIF4	1
T 2	STRAGE2	LSTIF5	2

5011 ストレージ装置 I/F テーブル

【図 13】

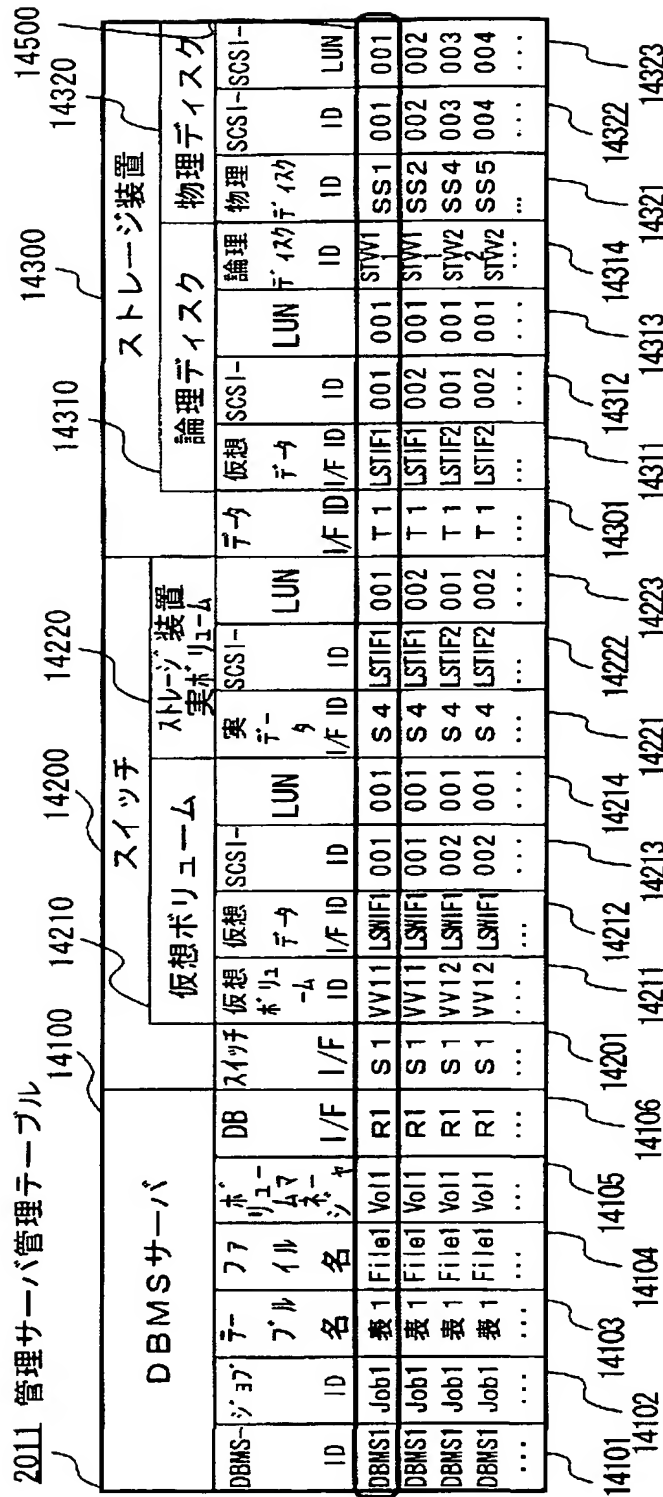
図 13

論理ディスク			物理ディスク			
仮想データ/F ID	SCSI-ID	LUN	論理ディスクID	物理ディスクID	SCSI-ID	SCSI LUN
LSTIF1	1	1	STVV11	SS 1	1	1
LSTIF1	1	2	STVV11	SS 2	2	2
LSTIF2	2	1	STVV22	SS 4	3	3
LSTIF2	2	2	STVV22	SS 5	4	4
LSTIF3	3	1	STVV33	SS 6	5	5
LSTIF3	3	2	STVV33	SS 7	6	6
LSTIF4	1	1	STVV44	SS 8	7	7
LSTIF4	1	2	STVV44	SS 9	8	8
LSTIF5	2	1	STVV55	SS 10	9	9
LSTIF5	2	2	STVV55	SS 11	10	10

5012 ストレージ装置管理テーブル

【図 14】

図 14



【図 1 5】

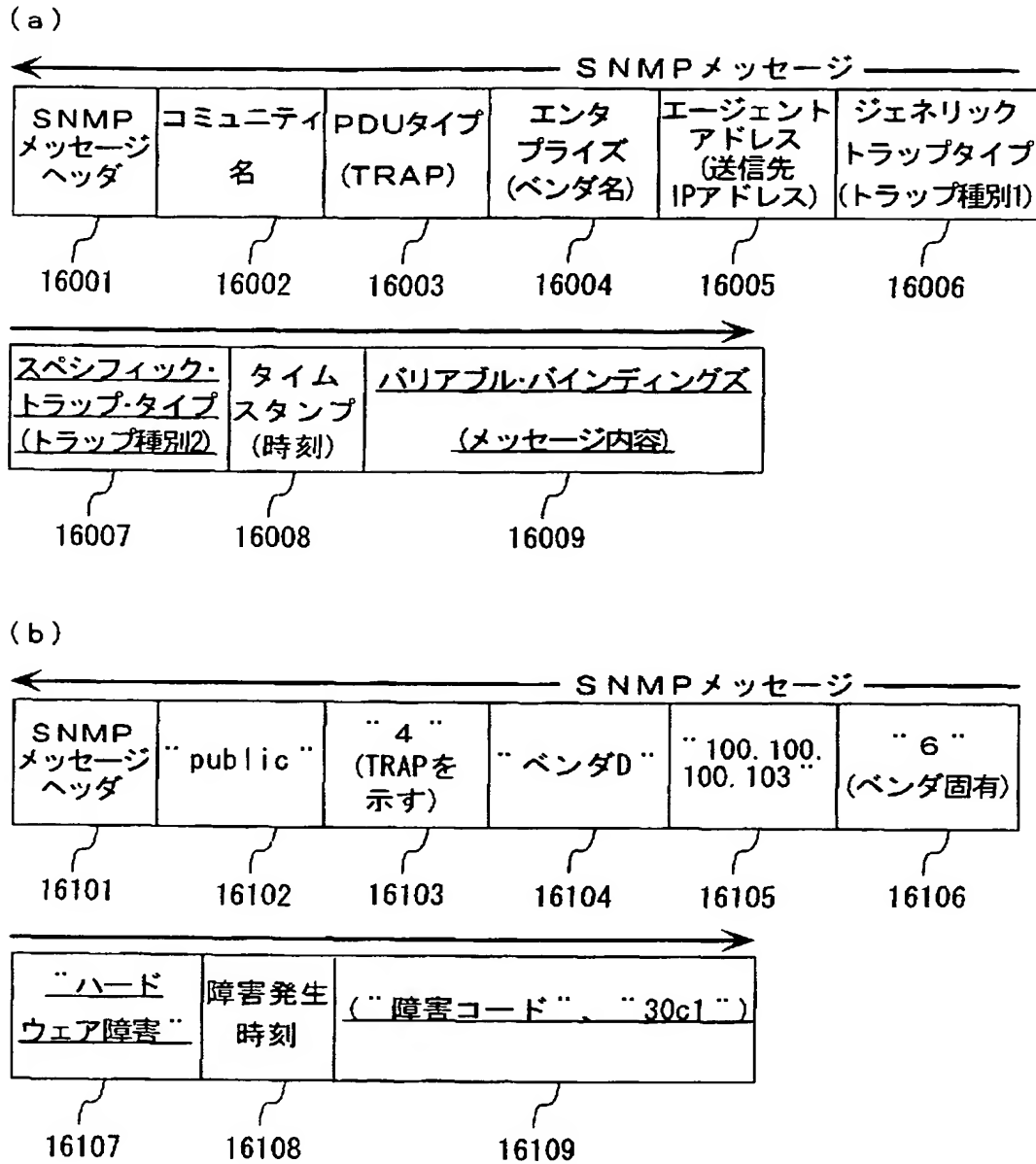
図 1 5

15010 障害 コード	15020 障害/復旧部位	15030 識別子	15040 障害管 理 I D	
00a1	実ボリ्यूーム復旧	S S 1	F 3031	15101
00a2	実ボリ्यूーム復旧	S S 2	F 3032	15102
...	...	...	...	
30c1	実ボリ्यूーム障害	S S 1	F 3031	15103
30c2	実ボリ्यूーム障害	S S 2	F 3032	15104
...	...	...	...	

2012 イベント辞書

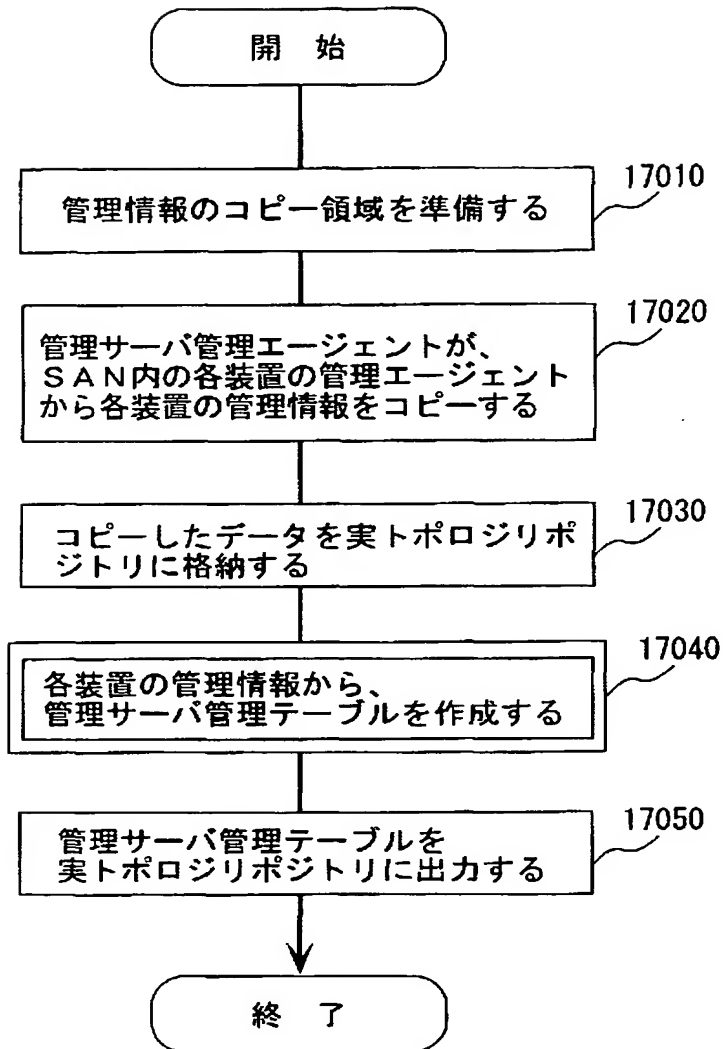
【図 16】

図 16



【図 17】

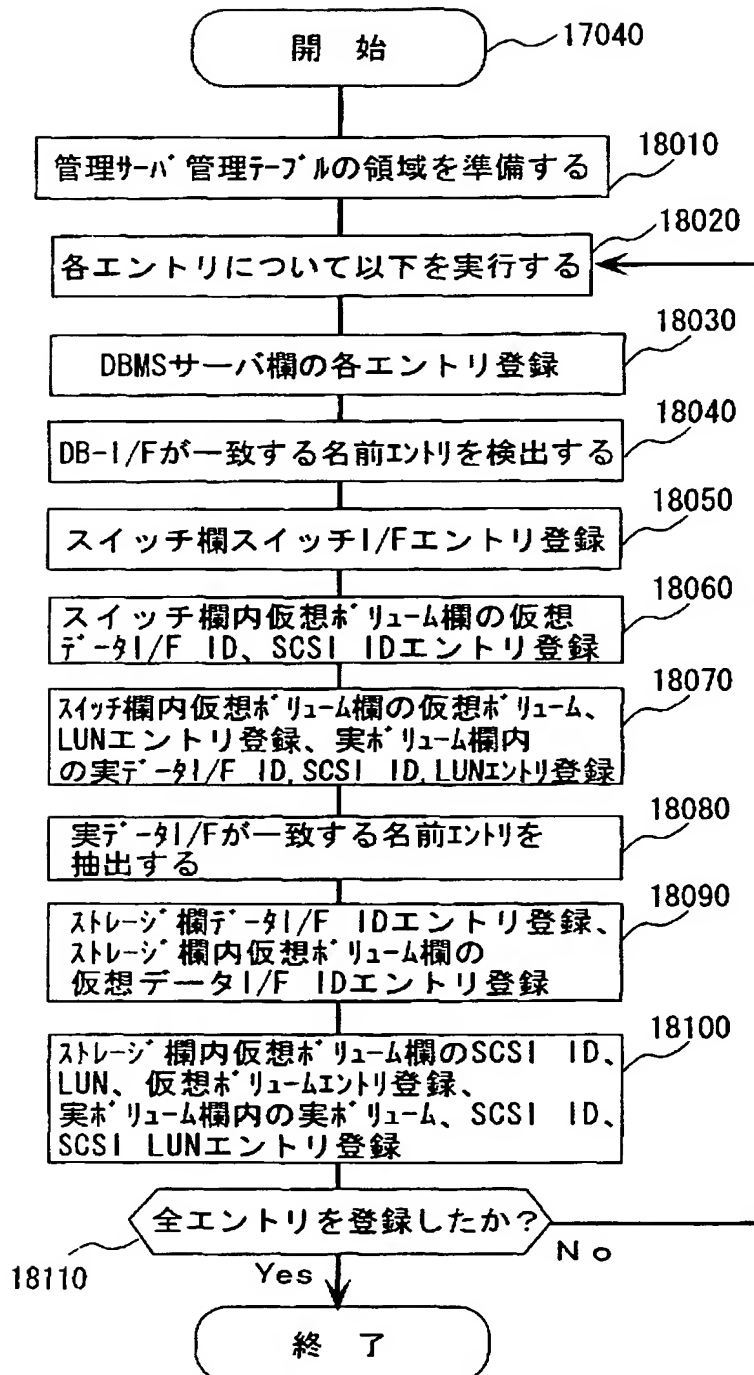
図 17





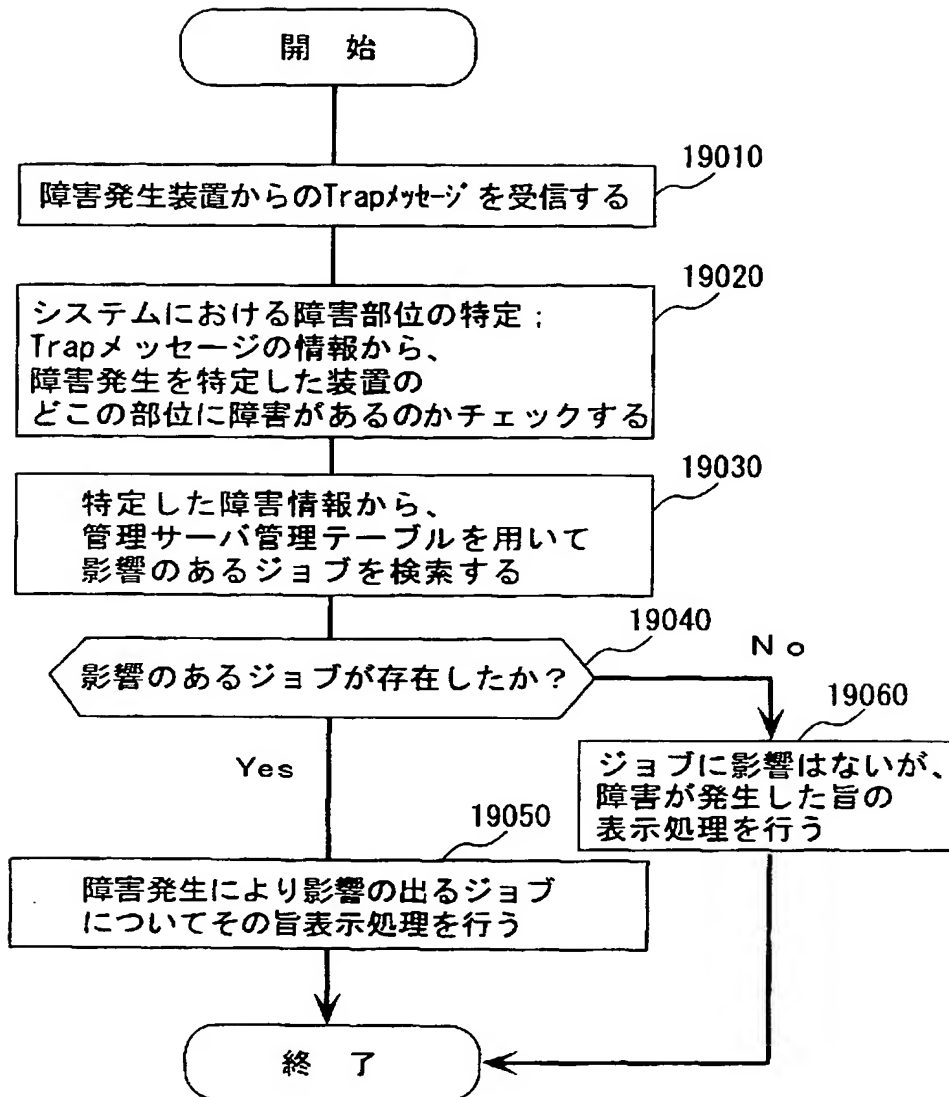
【図 18】

図 18



【図 19】

図 19



【図 20】

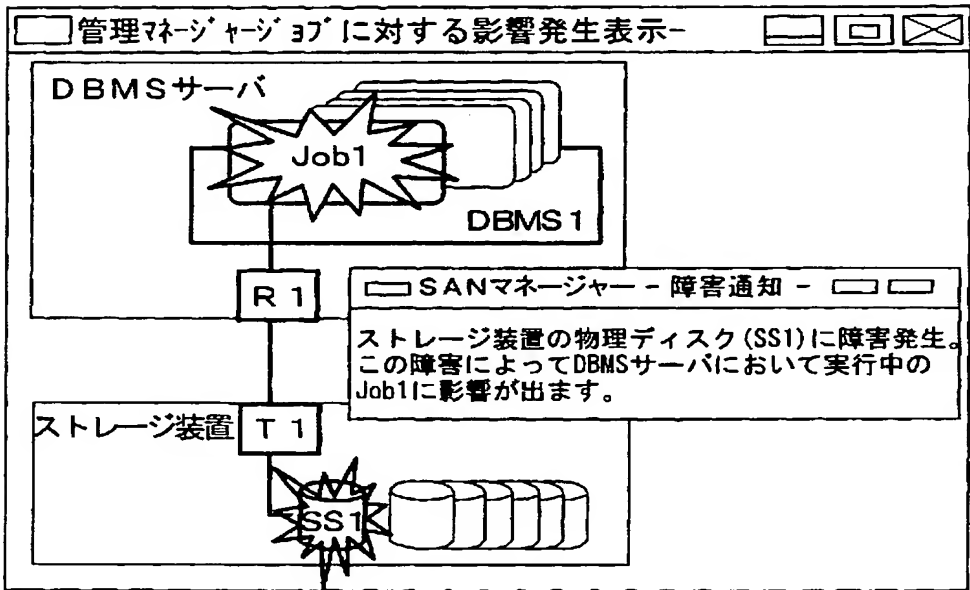
図 20

20000 管理 障害 ID	20050 DBMSサーバ					20200 ストレージ			
	DBMS ID	ジョブ ID	テーブル 名	ファイル 名	DB I/F	データ I/F ID	実ボリュームID	SCSI ID	SCSI LUN
	F3031	DBMS 1	Job1	表1	File1	R1	T1	SS1	1

20051 20110 20120 20130 20140 20150 20210 20220 20230 20240 20500

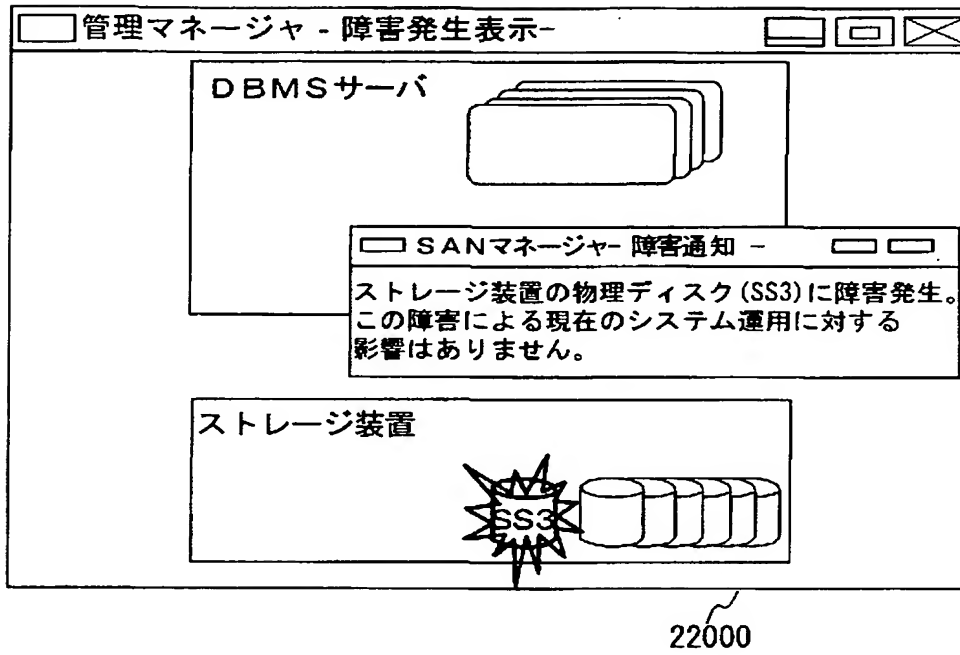
【図 21】

図 21



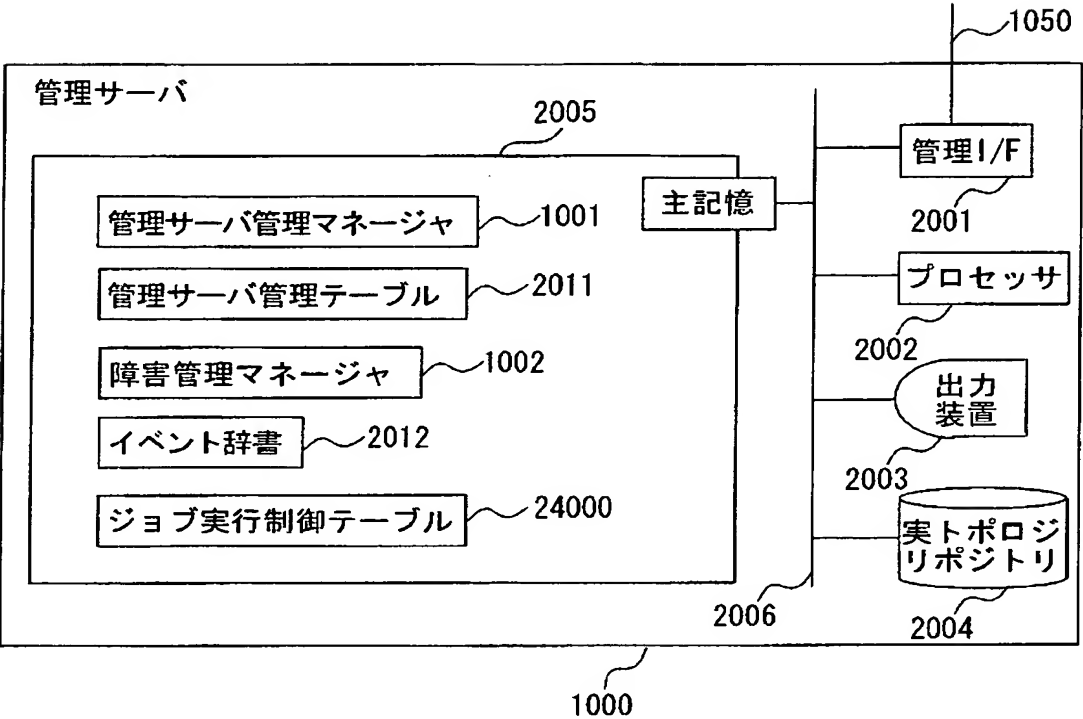
【図 22】

図 22



【図 2 3】

図 2 3



【図 2 4】

図 2 4

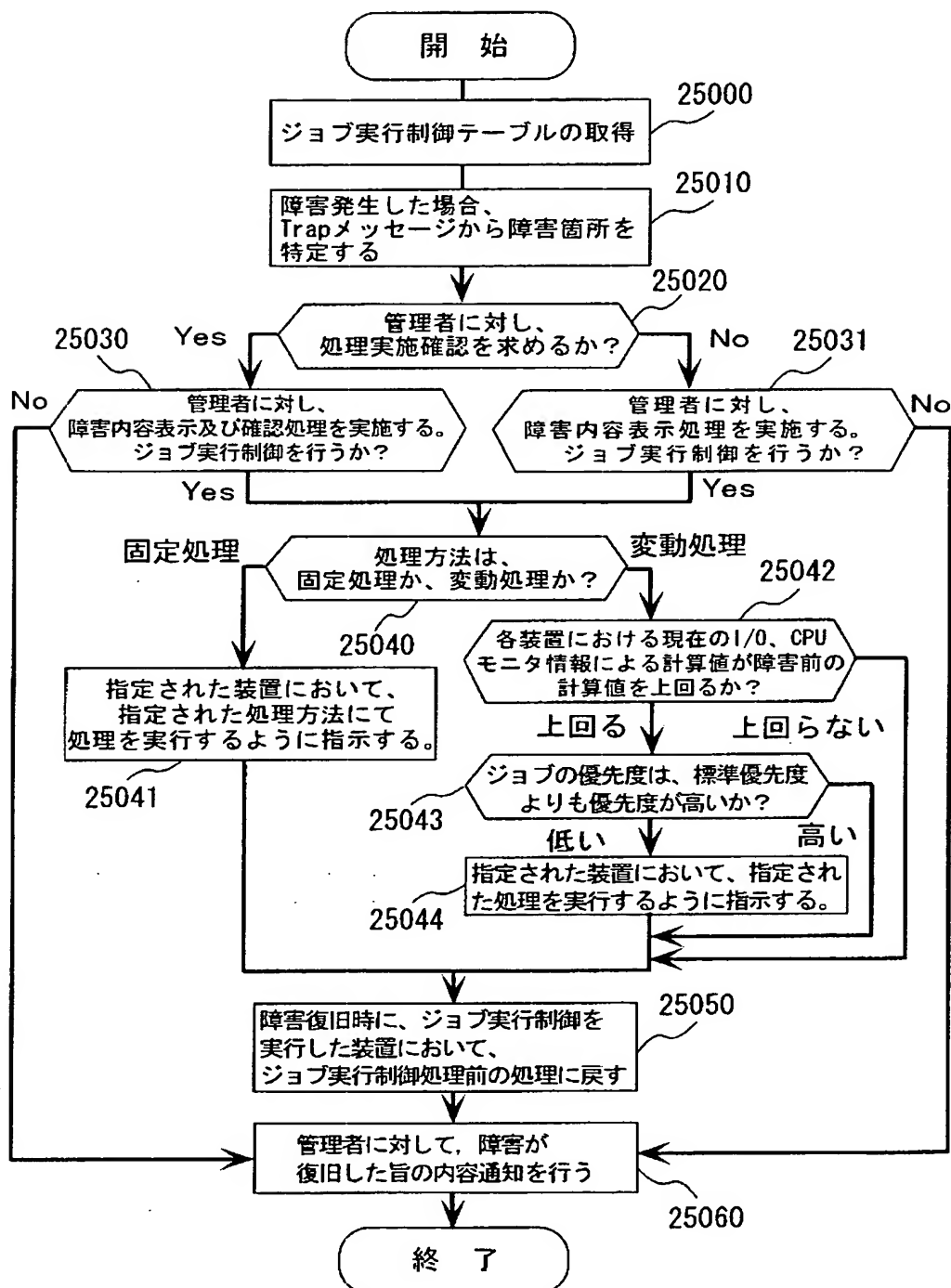
24000 ジョブ実行制御テーブル

DBMS ID	ジョブ ID	実施 装置	優先 度	処理 方法	処理 内容	確認 処理	自動 実施
DBMS 1	Job1	DBMS	1	固定	閉塞 処理	必要	--
DBMS 1	Job2	DBMS	2	変動	休止	不要	実施
...	...	...	...	...	...	...	...

24101 24102 2410 24020 24030 24040 24050 24060 24070 24080

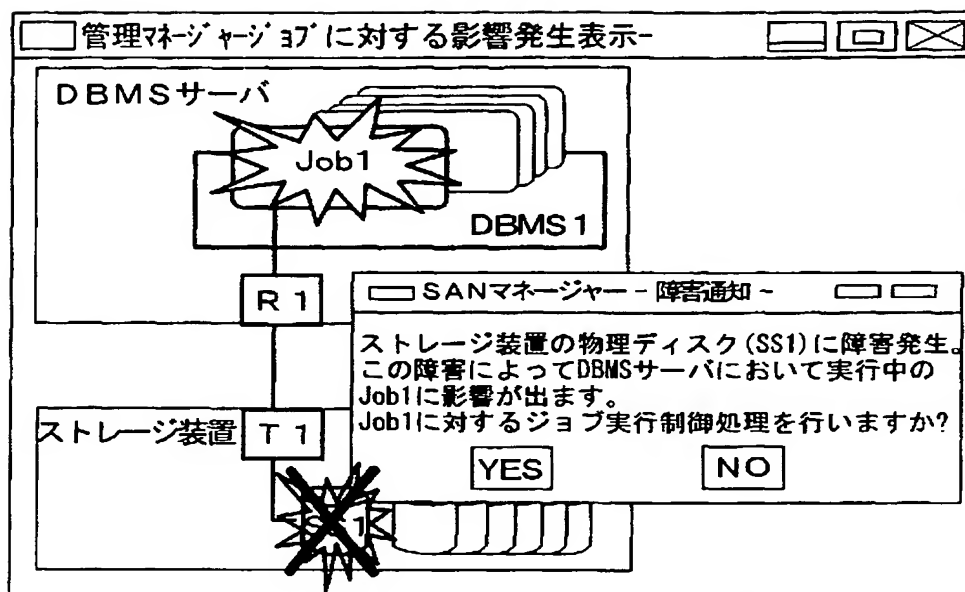
【図 25】

図 25



【図 26】

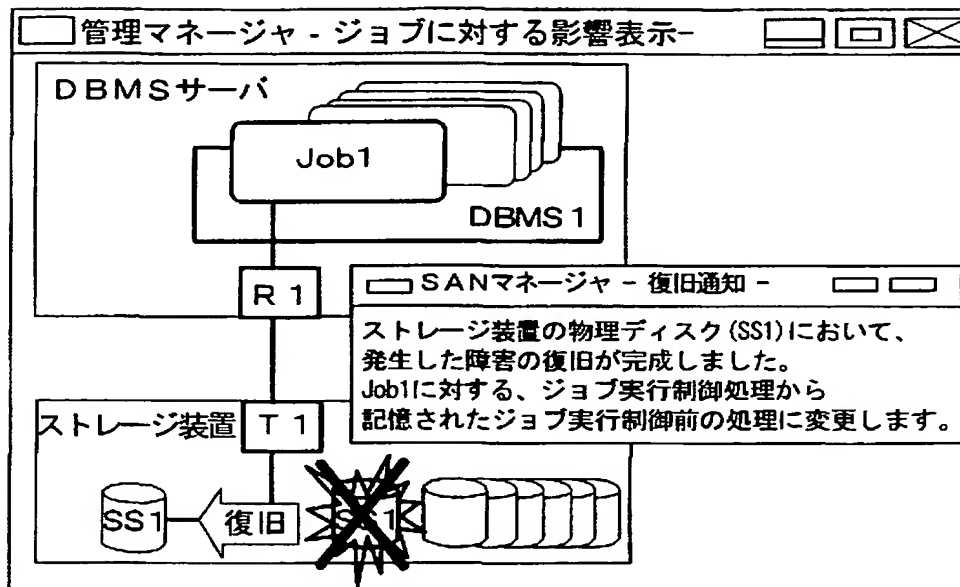
図 26



26000

【図 27】

図 27



27000



【書類名】 要約書

【要約】

【課題】 ストレージ系のいずれかの装置、機構に障害が発生したとき、この障害によって影響を受けるジョブを特定し、これらジョブの実行制御をする。

【解決手段】 DBMSサーバ装置1010、仮想化スイッチ装置1020及びストレージ装置1030の各装置は、ジョブがアクセスする特定のDBテーブルに始まり、このテーブルを格納する論理的なボリュームを経てボリューム上のデータを分散して格納する物理ディスク装置に至るまでのデータマッピング情報の一部情報を保持する。管理サーバ装置1000は、各装置からその一部情報を取得し、ジョブごとにデータマッピング情報を集約する。データマッピング経路上のいずれかの装置、機構に障害が発生したとき、上記データマッピング情報を参照して障害によって影響を受けるジョブを特定する。また管理サーバ装置1000は、これらのジョブの実行制御をする。

【選択図】 図 1

特願 2 0 0 3 - 0 9 0 5 1 8

出 願 人 履 歴 情 報

識別番号

[ 0 0 0 0 0 5 1 0 8 ]

1. 変更年月日

1 9 9 0 年 8 月 3 1 日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台 4 丁目 6 番地

氏 名

株式会社日立製作所